

Enter the Landscape Matrix: Effects of Anthropogenic Landscapes and Land Management on the Floral Quality of Tallgrass Prairies in Illinois, USA

Andy Sima



Enter the Landscape Matrix: Effects of Anthropogenic Landscapes and Land Management on the Floral Quality of Tallgrass Prairies in Illinois, USA

Andy Sima

Institutionen för naturgeografi/Department of Physical Geography
GE9011 Degree Project in Physical Geography and Quaternary Geology (60 HE Credits)
Series: NKA 412
Master's Programme in Landscape Ecology (120 HE Credits)
Spring term 2025
Supervisor: Anna Treydte
Examiner: Sara Cousins
Assessor: Jessica Lindgren
© Author: Andy Sima (Andrew David Sima)
Cover: Andy Sima



Stockholms
universitet

Abstract

North American tallgrass prairies, grassland ecosystems defined by their adaptations to fire and grazing, are one of the most endangered ecosystems on earth. In the American state of Illinois, less than 1% of original tallgrass prairie remains. Today, original remnant prairies and replanted restoration prairies are surrounded by a sea of anthropogenic landscape matrices, most predominantly agriculture and suburban development. Remnant and restored prairies both require active management, including burning, mowing, herbicide, seeding, hand weeding, and other management techniques. They are managed by a large number of disparate organizations with a wide range of socioeconomic backgrounds. Little is known about how anthropogenic landscape matrices and external socioeconomic variables, including land management and ownership, affect prairie floral diversity and quality. I sampled thirty prairies across northern and central Illinois through August and September 2024. At each prairie, I inventoried herbaceous species in ten 1 m² quadrats. Calculating the native species richness, floristic quality index (FQI), and other floral quality measures, I compared the diversity and quality of prairie zones against an array of external variables. These external variables included landscape matrix metrics, derived from the American National Land Cover Database (NLCD), and management practices and socioeconomic variables, collected from land manager responses to an email survey. I found that prairies in rural landscapes generally have higher species diversity, FQI, and Shannon diversity values, suggesting that rural prairies are of a higher quality than prairies in suburban landscapes. I also determined that sites that employ hand weeding as a management tool, as well as sites with greater inputs of human labor, are generally correlated to higher floral diversity and higher floral quality in these prairies. I conclude that if organizations in Illinois are to maintain the few tallgrass prairies that are left today, then management plans should take the effects of the surrounding anthropogenic landscape matrix into account. Organizations should also explore increasing high-labor management techniques, such as hand weeding, if their organization's management resources allow for its implementation in their prairies.

Keywords

Tallgrass prairies, grasslands, landscape matrices, biodiversity, land management, restoration, preserved ecosystems, anthropogenic landscapes, prescribed burning, grassland management, management techniques, floristic quality index, species diversity, floral quality, coefficient of conservatism, survey data, species coverage, Shannon diversity

Contents

1	Introduction.....	1
1.1	Prairies in the United States of America	1
1.2	Reduction in Tallgrass Prairies	2
1.3	Remnants and Restorations.....	4
1.4	Prairie Management	4
1.5	Threats to Prairies.....	5
1.6	Study Aim.....	6
2	Methodology	8
2.1	Primary Data Sources.....	8
2.1.1	Prairie Definitions and Sampling Terminology	8
2.1.2	Site Selection	10
2.1.3	Data Collection	12
2.1.4	Statistics for Vegetation Data	13
2.2	Secondary Data Sources	14
2.2.1	Land Cover Datasets and GIS Analysis	14
2.2.2	Management Survey Data.....	16
2.3	Statistical Analysis	17
3	Results	19
3.1	Anthropogenic Landscape Matrix.....	19
3.2	Management	22
3.3	Age, Budget, and Other Socioeconomic Factors	25
3.3.1	Age of Restoration	25
3.3.2	Organizational Funding	26
3.3.3	Ownership and Legal Designation	27
3.3.4	Area.....	27
3.4	Multivariate Modelling and Sub-Grouping	27
4	Discussion	29
4.1	Anthropogenic Landscape Matrix.....	29
4.2	Management and Socioeconomic Considerations	31
4.3	Multivariate Analyses	32
4.4	Future Directions	32
4.5	Synthesis and Conclusions	33
	Acknowledgements and Attributions	35
	References.....	36
	Appendix Information.....	43

1 Introduction

Grasslands are one of the most extensive ecosystems in the world; in the United States of America alone, the original extent of the grass-dominated Great Plains ecoregion covered over 350 million hectares, totaling between 15% and 30% of the continental United States (Clark 2011). Because of its size, there are wide spectrums of precipitation and temperature, creating a vast array of ecosystem subtypes and floral communities. One grassland subtype is the tallgrass prairie, which makes up the eastern side of the Great Plains (Figure 1, from Theshibboleth 2017). Prior to Euro-American colonization of the North American continent, some studies estimate that the United States component of the tallgrass prairie spanned over 60 million hectares, stretching from the Canadian border in the north to the Mexican border in the south, and as far east as Lake Michigan and as far west as central Nebraska (Steinauer and Collins 1996; Corbett 2004).

1.1 Prairies in the United States of America

Forming around 10,000 years ago after the retreat of the last glaciation, the tallgrass prairie is relatively mesic in comparison to other parts of the Great Plains (Steinauer and Collins 1996; Anderson 2006). However, like the rest of the Great Plains, tallgrass prairies are still defined by their climatic extremes. The tallgrass prairies receive between 600 mm and 1000 mm of rain per year, typically during the warm growing season, though droughts are common (Steinauer and Collins 1996; Knapp et al. 2006). Temperatures can range from 45° Celsius to -35° Celsius (Steinauer and Collins 1996). The bulk of biomass is made up of an association of tall (sometimes > 2 m) C4 grasses: most dominantly, big bluestem (*Andropogon gerardii*) and Indiangrass (*Sorghastrum nutans*), along with shorter grasses like little bluestem (*Schizachyrium scoparium*) (Weaver and Fitzpatrick 1932; Clark 2011). Despite the bulk of biomass being made up by only a few species, forbs account for a large percentage of prairie diversity. However, each prairie's exact composition is highly variable (Bragg and Steuter 1996).

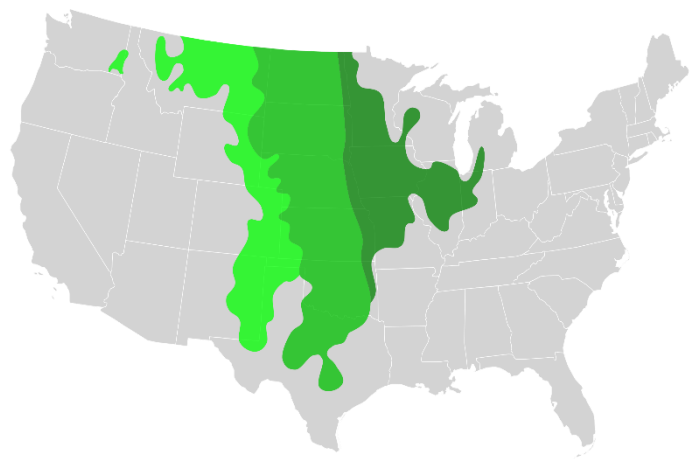


Figure 1. A map demonstrating the approximate United States extent of the central grassland ecosystems, otherwise known as the Great Plains (Theshibboleth 2017). Colors denote the approximate areas distinguished by specific grass species associations. The darkest green eastern portion is the tallgrass prairie, where grasses can grow >2 m tall. The middle section represents the mixed grass prairies, and the light green western section represents the shortgrass prairie (grass typically <1 m). This paper's scope is limited to tallgrass prairie within Illinois, representing some of the easternmost extent of the tallgrass prairie and the Great Plains.

Besides being defined by climatic extremes, tallgrass prairies are adapted to regular extreme disturbances. Most prominently, tallgrass prairies historically saw wildfires

every two to five years, which would burn vast swaths of land (Allen and Palmer 2011). Additionally, tallgrass prairies were home to the American Bison (*Bison bison*), a large grazing mammal with herds that could number in the thousands or higher (Knapp et al. 1999). Both of these regular disturbances served several functional roles within the ecosystem, including recycling nutrients, reducing cover, dispersing seeds, preventing woody encroachment, and creating greater habitat heterogeneity (Steinauer and Collins 1996). It is likely that both disturbances were managed by indigenous peoples for thousands of years prior to Euro-American colonization (Anderson 2006).

Tallgrass prairies offer unique qualities and ecosystem services, especially within the context of the Illinois landscape matrix. Soil generation is one of the most obvious, as prairie soil is responsible for the fertility of farmland in Illinois (Knapp and Smith 2001; Baer et al. 2003; Mackelprang et al. 2018). However, tallgrass prairies also serve to prevent soil erosion, capture, store, and clean scarce rainwater, sequester high amounts of atmospheric carbon dioxide, provide pollinators with natural habitat, offer resources to migratory birds, provide habitat to game and rangeland species such as deer and bison, and offer a wholly unique aesthetic and ecological experience (Steinauer and Collins 1996; Williams and Diebel 1996; Ampleman, Crawford, and Fike 2014). The majority of the prairie's ecosystem services have been lost since Euro-American colonization due to the massive habitat loss at all ecosystem scales (Samson and Knopf 1994; Suding 2011; Chalcraft et al. 2004; Koerner et al. 2016). Although land conversion has largely ceased, climate change models predict that these reductions will only get worse (Craine et al. 2011).

1.2 Reduction in Tallgrass Prairies

While Euro-American colonization of the Great Plains has drastically altered all Great Plains landscapes, tallgrass prairies were hit the hardest, and now represent one of North America's most endangered ecosystems (Rice et al. 2018). The tallgrass prairie ecosystem covered nearly 9 million hectares in the American state of Illinois alone (Corbett 2004). This historically represented some of the easternmost extent of the Great Plains. However, by the beginning of the 20th century, around 99.9% of Illinois' original tallgrass prairie had been extirpated (Samson and Knopf 1994). This dramatic reduction is due entirely to human action, and the majority of tallgrass prairie was converted to agriculture. Because of the tallgrass prairie's exceptionally fertile soil, its conversion was deemed highly desirable for American expansionism (Weaver 1927; Petipas 2024). To this day, farmland supported by tallgrass prairie soil is a cornerstone of North American agricultural output (Fowler 2024; United States Department of Agriculture 2024).

The complete and total alteration of the tallgrass prairie was carried out in several ways. For many prairie scholars, the development of the steel plow by John Deere in 1837 is considered the death knell of the tallgrass prairie. Crucially, it was one of the first farming tools that allowed Euro-American colonizers to quickly and easily break up the tough prairie sod and begin farming (Smith 1992). Following that, prairies were further degraded through the suppression of natural wildfires and the extirpation of all American Bison from their original Illinois ranges (Knapp et al. 1999). The eradication and removal of indigenous peoples from their ancestral prairie homelands further advanced prairie loss (Anderson 2006). Along with urban development, these mechanisms have resulted in a near-complete destruction of tallgrass prairie, replacing it with a modern monoculture of corn or soy that covers the vast majority of the American Midwest, and most of Illinois (Yu and Lu 2018). Illinois's agricultural space covers around 11 million hectares (Illinois Department of Agriculture 2025).

Remaining tracts of original prairie survived largely due to the presence of unfarmed railroad verges, settler cemeteries, or soil that was too wet or too rocky to properly farm

(Steinauer and Collins 1996; Robertson, Anderson, and Schwartz 1997). In today's tallgrass prairie conservation, there is a dichotomy between "remnant" prairies and "restored" prairies. Remnant prairies are parcels of original tallgrass prairie that were never plowed under and have remained tallgrass prairie since prior to Euro-American colonization (Robertson, Anderson, and Schwartz 1997; Anderson 2006; Bach and Kleiman 2021). Fewer than 1,000 hectares of remnant prairie remain in Illinois (Samson and Knopf 1994). Restored prairies, for the purposes of this paper, are prairies that were removed in their entirety and converted to another land cover type (typically agriculture), but have since been restored to tallgrass prairie at some point in the recent past. Most of the time, these restorations are undertaken solely for conservation purposes.

Today, most prairies, both remnant and restoration, are small and hedged in on some or all sides by woodland, agriculture (typically corn), suburban development, or industry (typically factories or coal energy production in Illinois). This fragmentation would have been highly unusual prior to European colonization, when prairies would stretch unbroken for hundreds of kilometers. The largest prairies in Illinois today are typically also internally fragmented by smaller railways, service roads, pipelines, abandoned industrial infrastructure, or electrical wire (Figure 2, cover image).



Figure 2. An array of images from field data collection. Clockwise from top, images are from Goose Lake Prairie, Shoe Factory Road, Goose Lake Prairie, and Nachusa Clear Creek 2006. Images have been selected to give readers a general idea of the appearance of Illinois tallgrass prairie while also demonstrating certain landscape characteristics typical of prairie preserves in Illinois, such as their high levels of fragmentation, edge encroachment by woody species, and the inclusion of human infrastructure.

Cover Image. Images on the cover of this report are, from top left, clockwise: New England Aster (*Symphyotrichum novae-angliae*), Nachusa Thelma Carpenter prairie with a covering of showy goldenrod (*Solidago speciosa*) and other forbs, and Weston Cemetery Prairie. These cover images similarly serve to demonstrate examples of Illinois prairie, including the state of the rural landscape matrix surrounding most remaining prairie parcels.

1.3 Remnants and Restorations

The goal of tallgrass prairie restorations is generally to mimic remnant prairies and provide similar habitat and ecosystem services. However, it is well established that these two types of prairies have significant differences. Typically, prairie restorations have reduced species diversity than prairie remnants. While this may be due to restoration seed mixes lacking enough species, it is also theorized that this may be due to restorations lacking enough functional plant groups in close proximity to maintain diversity (Polley, Derner, and Wilsey 2005). Even when remnants and restorations have comparable species richness, remnant prairies have much higher phylogenetic diversity, and are generally more temporally stable (Barak et al. 2017; Sluis, Bowles, and Jones 2018).

Remnant and restoration prairies also tend to respond differently to threats and disturbances. Remnant prairies are less affected by irregular disturbances such as drought; restored prairies, even those over fifty years old, tend to be harder hit by drought (Allison 2008). Furthermore, remnant ecosystems are more resistant to invasion by nonnative species (Rowe 2013).

The vast majority of tallgrass prairies in Illinois today are restorations, sometimes as wholesale restorations and sometimes as buffers for a remnant prairie core (Rowe 2013). Most typically, restoration has been undertaken on agricultural fields within the last 50 years (Taft, Hauser, and Robertson 2006). Because these ecosystems are so fragmented, yet require a wide variety of disturbances and ecosystem functions at different spatial scales, they act as islands among seas of suburban housing and industrial agriculture (Polley, Derner, and Wilsey 2005). In all but the largest prairies (none of which are in Illinois), biodiversity and ecosystem quality cannot reliably be considered as self-sustaining and require active management (Koper, Mozel, and Henderson 2010).

1.4 Prairie Management

Today, nearly all prairie restorations, and the majority of prairie remnants, are highly-managed conservation spaces. Management practices aim to increase the biodiversity and quality of native species, sometimes with the specific goal of protection or reintroduction of threatened species (Steinauer and Collins 1996; Campbell and Hooymans 2016; Bach and Kleiman 2021). Management may be carried out on privately- or publicly-owned land and may be carried out by any number of managing organizations, including private non-profits or municipal, state, or federal government entities (Williams 2005). There are significant social, political, and economic differences between managing organizations, and often there is little-to-no organizational, administrative, or bureaucratic overlap between disparate conservation units (Williams 2005; Moilanen and Arponen 2011). All management activities require significant resource expenditure of both time and money, but the resources available to any given organization can vary enormously (Collins 2000).

With effort and labor, it is possible to maintain diverse tallgrass prairie communities through active management (Bach and Kleiman 2021). The most common management tools for Illinois tallgrass prairie conservationists are controlled burns and mowing, along with native seeding, application of herbicides, and hand-removal of invasive species (Harmon-Threatt and Chin 2016). Management efforts attempt to increase diversity by replicating natural disturbance forces and dealing with modern threats, but managing organizations may not have the capacity to replicate the complex pre-settlement disturbance regimes of tallgrass prairies (Steinauer and Collins 1996). Many land managers state that labor and funding are the biggest limiting factors for quality management (Rowe 2010).

The responses of managed prairies to individual management techniques are well established, though research has primarily focused on fire and grazing, while the combinations of various management techniques are less well understood (Samson et al. 1996, Wagle and Gowda 2018). Infrequent (every 2-5 years) burning can increase forb cover and diversity while maintaining grass extent; well-managed burning can also prevent the spread of woody species, which tend to dominate prairies in the absence of disturbance (Reemts and Belcher 2022). Grazing is also believed to be a crucial disturbance that would have been historically typical of tallgrass prairies; it can increase soil nitrogen and, like burning, increases sunlight at the ground level (Steinauer and Collins 1996). However, both burning and grazing tend to select for certain subsets of species; ecosystems with both disturbances tend to have higher diversity overall (Larson et al. 2020). Unfortunately, the reintroduction of bison is relatively uncommon in Illinois. No sites in this study utilized grazing, and mowing was the closest comparable management technique. Mowing simulates grazing in that it removes top-cover of plants, recycles nutrients, and increases light levels, but differs from grazing in distinct ways (Steinauer and Collins 1996; Römermann et al. 2009). Mowing, unlike grazing, is homogenous and non-selective. As most sites lack the funds or space to graze bison, mowing is the typical alternative (Steinauer and Collins 1996; Römermann et al. 2009). However, there is a knowledge gap regarding combinations of management tactics, certain understudied tactics, and how the socioeconomic makeup of a site may influence its management and, therefore, its prairie.

1.5 Threats to Prairies

Some of the long-term mechanics of prairie loss continue to be felt today, such as the absence of keystone faunal species and their associated functional diversity, largescale fragmentation of existing prairies and the complex detrimental effects of habitat fragmentation, and the continued suppression of fire and other natural disturbance regimes in the name of the surrounding landscape matrix (Steinauer and Collins 1996, Bragg and Steuter 1996).

New threats are also felt in the increasing competition by exotic introduced species and the long-term climatic changes brought on by worldwide climate change (Avolio 2021). Most managers find that invasive species are one of the greatest threats to prairie conservation, but climate change is likely to exacerbate existing threats (Rowe 2010; Avolio 2021). As the global climate changes, land managers may have to adapt management techniques to account for reduced precipitation and increased temperatures (Wagle et al. 2019). There is also still the constant threat of prairies being destroyed to make way for industrial or suburban development, such as in the case of the Bell Bowl Prairie's destruction for an airport service road (Wetli 2023).

However, the scale and intensity of these threats can differ greatly depending on a prairie's location (Jules and Shahani 2003). Especially in northern Illinois, there is another dichotomy beyond that of remnant and restoration. Besides its huge agricultural landscape, Illinois is home to the Chicago metropolitan area; a vast, sprawling patchwork of urban, peri-urban, suburban, and natural spaces. The Chicago metropolitan area, defined as the continuous developed/residential space extending beyond city boundaries, is home to over nine million people and covers an area greater than 2.5 million hectares (Reinhold 2025).

Speckled within these two suburban and agricultural landscapes are the last remaining vestiges of tallgrass prairie. For the purposes of this study, "suburban" prairies may be surrounded by any range of human development, including residential neighborhoods, golf courses, parking lots, factories, energy production infrastructure, transportation infrastructure, and more. "Rural" or "agricultural" prairies generally have a more consistent surrounding landscape matrix, typically being enveloped by farmland

dedicated to corn or soy crops. This farmland comes with the associated infrastructure of drainage ditches, irrigation and farming machinery, roads, and railways. These matrices and their effects may also vary over time (Jules and Shahani 2003)

The effects that these disparate landscapes can have on tallgrass prairies are not well understood. One study suggests that edge effects (such as increased presence of invasive species, anthropogenic disturbances, including soil erosion, herbicide drift, and agricultural pollution runoff) play a greater role on prairie community structure than either landscape matrix or prairie area (Koper, Mozel, and Henderson 2010). They suggest that edge effects explain negative trends in tallgrass prairie diversity (Koper, Mozel, and Henderson 2010). However, other studies found that an increasing ratio of anthropogenic landscape matrix against patch size leads to reduced species richness, and prairies surrounded by natural environment are more likely to support higher levels of ecosystem functionality and quality than those surrounded by anthropogenic landscape matrix (Bascompte and Rodríguez 2001; Esparrago and Kricsfalussy 2015; Zirbel et al. 2019).

There is evidence that prairies adjacent to roads or abandoned land have a higher invasion rate than prairies adjacent to agricultural land (Rowe 2013). One study found that landscape had little impact at all on tallgrass prairie, but did not include a full gradient of anthropogenic landscape types (Grman, Bassett, and Brudvig 2013). However, beyond this there are few studies of anthropogenic landscape matrices, and fewer specific to North American Tallgrass prairies. While at least one later study found no change in intraspecific diversity among plants along an urbanization gradient, and there is some suggestion that urban prairies can reach the same levels of quality and functionality of rural prairies, a study carried out in Australia demonstrated that urban matrix grasslands had higher extinction rates than agricultural matrix grasslands (Kim and Noto 2025; Coleman and Wait 2022; Williams et al. 2005). A French study demonstrated changes in plant functional traits along urban-rural gradients, including phenology and plant height, but did not measure for diversity or quality changes (Cochard et al. 2019). There is a knowledge gap regarding these anthropogenic landscape matrices.

1.6 Study Aim

As there are limited tallgrass prairies remaining in Illinois, and limited funding with which to conserve them, it is crucial that land managers and researchers understand how tallgrass prairies respond to differences in the landscape matrix and management. The effects of landscape variables in restoration are understudied generally, especially between types of anthropogenic matrices (Brudvig 2011). To the best of my knowledge, there are no previous studies that explore differences in floral quality between North American tallgrass prairies across gradients of anthropogenic landscape matrix types. Additionally, while there are many studies that explore the impacts of management, there are few studies that also explore broader socioeconomic factors such as organizational ownership and management finances. Furthermore, there are few studies that do any of this within the context of prairies within and around Chicago, one of the United States' most populous urban centers.

Therefore, it is my goal to contribute to prairie knowledge by comparing numerous unique tallgrass prairies, both within the bounds of the Chicago metropolitan area and within the larger agricultural landscape of Illinois. Specifically, I have three main questions:

1. How do indices of floristic biodiversity and quality differ within tallgrass prairies when measured across a gradient of anthropogenic landscape matrices, ranging from heavily developed suburban area to very rural agricultural areas?

2. How do these same prairies differ in terms of external management and socioeconomic variables, such as management intensity, management technique, age of restoration, staffing levels, and organizational resource expenditure? More importantly, do these differences affect floristic biodiversity and quality?
3. Do landscape and management variables affect prairie quality equally, or are their impacts different?

Answering these questions could lead to applications in conservation management and planning for long-term sustainability (Török et al. 2021). With my study, I aim to:

1. Help predict the potential quality of a future prairie restoration based on its surrounding landscape matrix.
2. Add to knowledge regarding where conservation funds should be allocated for the best possible conservation outcomes.
3. Help determine if there are external indicators for the success of a managed prairie, and how further measures can be implemented to ensure a restoration's success.

From this research, I hope to draw conclusions beneficial to the management of all tallgrass prairies, even beyond the ones characterized in this study.

2 Methodology

Utilized datasets can be broken down into two categories: primary data collection and secondary data sources. Primary data sources include in-field floral surveys and measurements derived from those surveys. I collected in-field data between August and September 2024. Secondary data sources include information gathered from outside datasets, namely the National Land Cover Database from the United States Geological Survey and metrics derived from the land manager survey responses (Figure 3).

2.1 Primary Data Sources

2.1.1 Prairie Definitions and Sampling Terminology

Though there are many different subcategories of grassland within the broader category of Illinois tallgrass prairie, this research focused on the most abundant subcategory within Illinois: the tallgrass mesic prairie (Samson and Knopf 1994; Steinauer and Collins 1996). These grasslands typically have soil high in organic material, with moderate pH and water retention capacity (Chapman and Brewer 2008). However, the most reliable gauge of prairie subtype is grass associations; especially in the field, mesic, wet, and sand prairies can best be distinguished by the present floral community (Table 1; Weaver and Fitzpatrick 1932; Steinauer and Collins 1996). However, it is of note that none of the presented grass associations are strictly exclusive to their respective prairie subtype.

My sampling design delineates between three types of study units: sites, zones, and plots (Figure 4). “Sites,” are herein defined as the official administrative boundaries of the land ownership unit that contains the prairie zone. Sites also include the entire unsampled space beyond the prairie zone, sometimes including vast non-prairie ecosystems and developed recreation facilities. Some sites contained only the zone and nothing else. “Prairie zones,” also referred to as “zones of interest,” are herein defined as the discrete sampling unit of a singular contiguous prairie, which I quantitatively

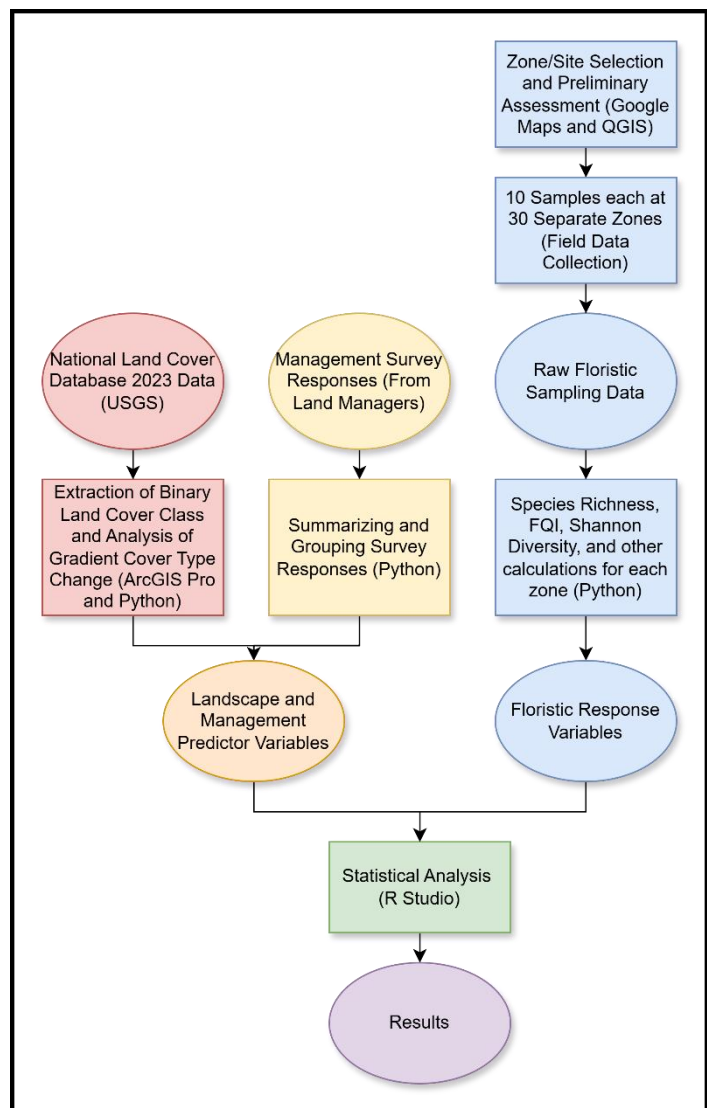


Figure 3. A methodological flowchart depicting the sources of data used in this research, as well as a summary of the tools and processes applied to the data. Colors are primarily for visual interest but roughly equate to the three utilized datasets (e.g., blue: primary in-field data and response variables, yellow: secondary management survey data, red: secondary NLCD data). Round shapes represent data while rectangular shapes represent a process applied to data.

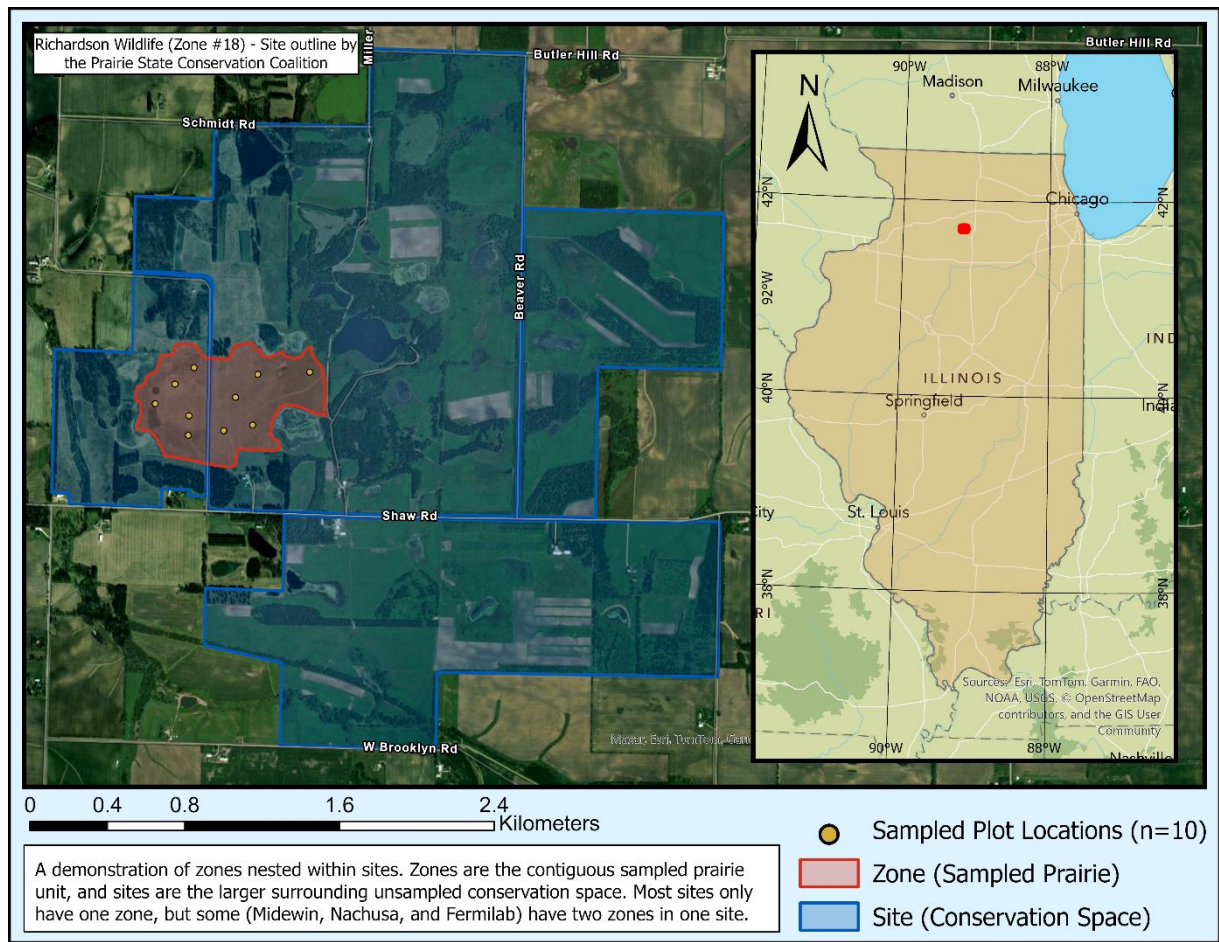


Figure 4. A demonstration of zones nested within sites. While each prairie zone is a contiguous unit of grassland, these zones were often contained within much larger conservation sites and were often surrounded by non-prairie ecosystem. Land ownership, budget, and other administrative variables were organized based on the site level. Zones typically have no legal or administrative designations in and of themselves and serve only to delineate the prairie fragment that was sampled for this study. Each zone contains ten sampled plots, which are the source of the floral surveys and coverage data that form the statistical basis of this study.

characterize through ten sampled plots. “Plots” are herein defined as the 1 m² survey quadrat that make up the basic sampling unit for this study’s floral survey component, and are the source of the in-field floral surveys.

Table 1. Information regarding the distinctions between the three primary types of prairies in Illinois. Soil data drawn primarily from Chapman and Brewer 2008, but also Batt 1996; Steinauer and Collins 1996. Grass association data drawn from Weaver and Fitzpatrick 1932; Batt 1996; Steinauer and Collins 1996. Values are given in ranges and (means). Standard soil textures and grass associations are given along with (less-common features).

* Denotes the primary research prairie subtype for this project.

Prairie Type	Soil pH	Soil Water Retention Capacity (%)	Soil Texture	Grass Associations
Sand	5.4-5.9 (5.5)	31-62 (43)	sandy loam (loamy sand, sand)	<i>Sporobolus heterolepsis</i> , <i>Schizachyrium scoparium</i> , (<i>A. Gerardii</i>)

Prairie Type	Soil pH	Soil Water Retention Capacity (%)	Soil Texture	Grass Associations
Mesic*	4.9-7.5 (6.2)	37-106 (64)	loam, sandy loam (silt loam)	<i>Andropogon gerardii</i> , <i>Sorghastrum nutans</i> (<i>Elymus canadensis</i> , <i>Panicum virgatum</i> , <i>Schizachyrium scoparium</i>)
Wet	5.5-7.7 (6.9)	43-170 (100)	loam (sandy loam, silt loam)	<i>Spartina pectinate</i> (<i>Juncus</i> sp., <i>Carex</i> sp.)

2.1.2 Site Selection

I developed a list of potential sites utilizing a variety of resources, primarily the Prairie Directory of North America, the Nature Preserves Directory, and the Prairie State Conservation Coalition’s dataset of Illinois protected lands (Adelman and Schwartz 2013; Illinois Department of Natural Resources 2024; Holman and the Prairie State Conservation Coalition 2024). I also discovered sites through pre-existing personal knowledge, advice from local experts, and online information from local governments. In order for a prairie to be eligible for study, it had to meet the following criteria:

- Within 200 km of Chicago city boundaries (Chicago Metropolitan Agency for Planning 2024)
- Within the state of Illinois
- Actively managed by natural resources professionals
- Mesic prairie (determined based on ecological site descriptions, plant lists, explorations in Google Earth (Google LLC 2022), direct land manager communication, and in-field validation)
- Assignable as either “rural” (surrounding landscape matrix greater than 50% farmland) or “suburban” (surrounding landscape matrix greater than 50% anthropogenic development and infrastructure) (Figure 5)

Through this process, I discovered around forty-five prairies that met these criteria, and settled on twenty-seven sites and thirty zones that I could sample within the available time frame. These sites were chosen by semi-random stratification so that they covered a large geographic area across northern and central Illinois while still meeting the need for an even number of rural and suburban sites (Figure 5).

As some sites were especially large and contained several unique, discontinuous prairies units, I broke three sites (Midewin, Nachusa, and Fermilab; see Appendix E) into two zones each, based on the guidance of land managers. Each zone within these sites was separated by several discontinuous breaks in the prairie, such as roads, non-prairie matrix, and waterways. In addition, each zone within these sites had distinct management regimes and histories within the larger context of the site’s management history. This left me with thirty sampleable prairie zones.

These thirty prairie zones were further divided based on surrounding landscape matrix, with fifteen rural and fifteen suburban each. While initial categorizations for landscape

matrix were visually estimated from satellite imagery, the matrix categories were later statistically validated to confirm the type of surrounding matrix (section 2.2.1).

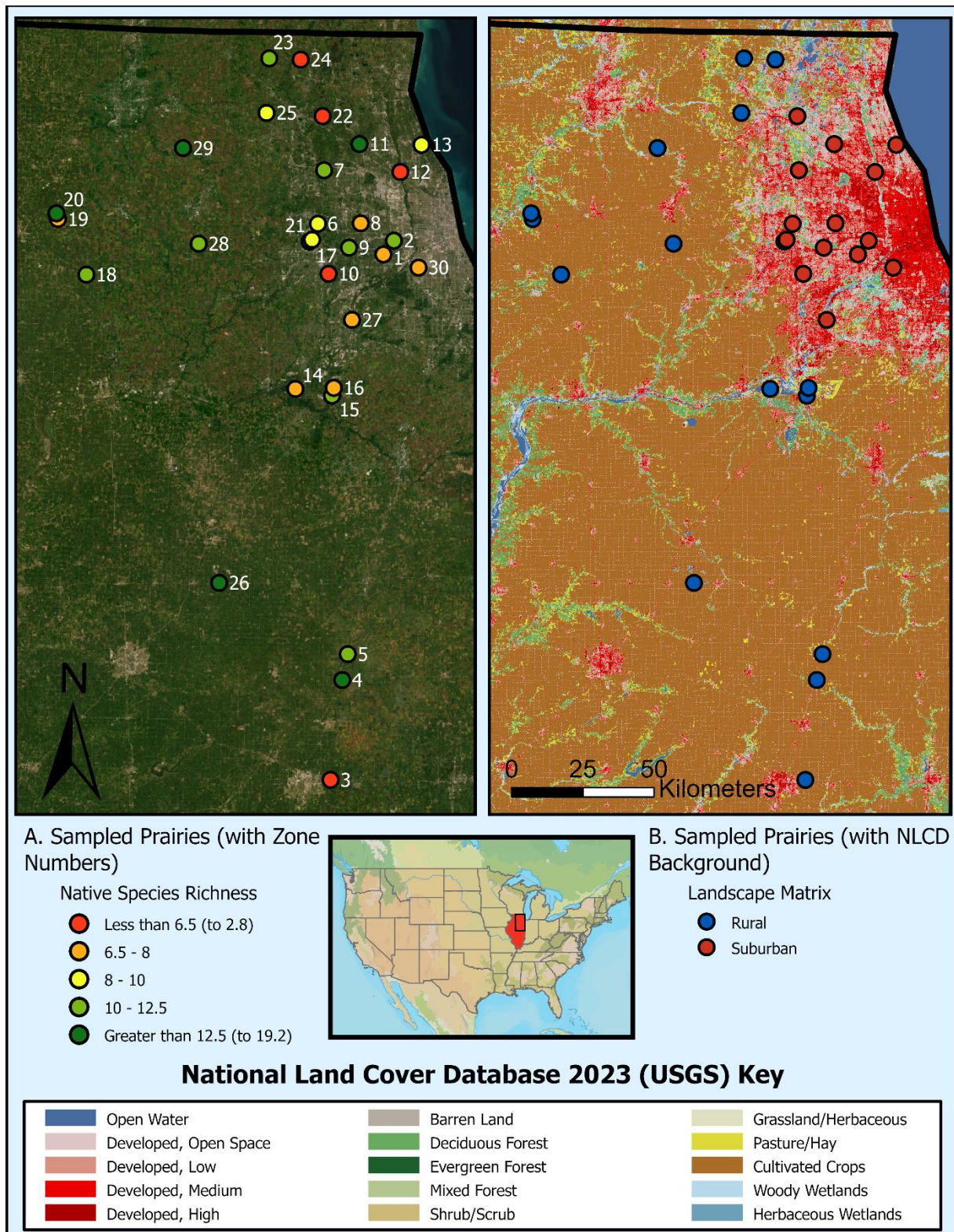


Figure 5. A map demonstrating the geographic distribution of all thirty sampled zones. The floral surveys collected from these thirty zones made up the backbone of the in-field primary dataset of this project. Zones were chosen from a list of potential prairies throughout northern and central Illinois, which was then refined to ensure that prairies were largely comparable in their ecology and management. Zones were selected to guarantee equal rural and suburban samples.

2.1.3 Data Collection

Once a relevant site and zone were chosen and the appropriate land use permissions acquired, I had to define the boundaries of the prairie zone. I used QGIS (QGIS Association 2024) to draw out tentative zone boundaries within the site, based on a combination of visual estimation of current and historic satellite imagery (Google LLC 2022), publicly available ecosystem maps, online ecosystem descriptions of sites, and direct land manager suggestions (where available). Efforts were made to ensure that each zone represented one prairie with a single, coherent management plan, later verified in-field.

I then used QGIS's random point tool and selected ten semi-stratified points from twenty randomly-distributed points within the zone. Exact sample point locations were finalized in-field so that, where possible, all plots were greater than 10 m (or as far as feasible) away from all of the following: prairie edge, established trails, permanent structures, standing or running water, woody vegetation with an average height greater than that of the surrounding prairie grass, and all other sample plots. In-field at the established plot locations, I took a floral survey within a marked 1 m² quadrat, ten per zone (Figure 6) (Guiden et al. 2021). Exact geographic location was recorded with the use of Avenza Maps (Avenza Systems Inc. 2024).



Figure 6. The survey quadrat as it was used in the field. Cross-grid wires were purely for visual coverage estimates, dividing the total area into eight blocks, each representing 12.5% of the total area; total area is 1 m². For exact dimensions, see Appendix Figure B17.

I assessed the biodiversity of each plot by identifying every plant within the quadrat at the species level (wherever possible; see appendix D), but did not assess leaf litter or dead plant matter. Identification was carried out with the help of Mohlenbrock's *Vascular Flora of Illinois* (Mohlenbrock 2014), as well as amateur field guides and neural network-based floral identification apps Seek and iNaturalist Classic (Chadde 2019; Kurz 2021; Van Horn et al. 2018; iNaturalist 2024 *iNaturalist*; iNaturalist 2024 *Seek*).

For each species, I assigned a modified Daubenmire canopy coverage class (Table 2). These Daubenmire canopy coverage classes borrow from previous prairie diversity studies (Abrams and Hulber 1987; Coulloudon et al. 1999; Higgins et al. 1996). Thus, species presence and coverage represent an approximation of the state of biodiversity and floristic quality for a given plot (Guiden et al. 2021).

Table 2. The breakdown of ranges and midpoints of the modified Daubenmire Coverage Classification system I utilized during the vegetation sampling portion of this research, primarily for the purpose of calculating the Shannon Diversity index

* Denotes no presence; absent species receive 0 automatically.

** Denotes a single individual instance of a species, unless plant is particularly large.

+ Denotes 2 or more individuals, but coverage still not exceeding 5% of total area.

Coverage Class	Coverage Percent	Midpoint
0*	0%	0
1**	0-1%	0.005
2 ⁺	1-5%	0.03
3	5-25%	0.15
4	25-50%	0.375
5	50-75%	0.625
6	75-95%	0.85
7	95-100%	0.975

2.1.4 Statistics for Vegetation Data

To determine a zone's biodiversity and quality, I calculated native species richness, percent of introduced species, native Floristic Quality Index (FQI; Taft, Hauser, and Robertson 2006), Adjusted Floristic Quality Index (AFQI; Rocchio 2007), and a modified form of the Shannon-Wiener Diversity Index for all species (hereafter called just "Shannon's diversity index"; Taft, Hauser, and Robertson 2006). I calculated these for all ten plots per zone and then averaged them for each variable to get floral diversity and quality variables for all 30 zones. This is excepting Shannon diversity, which requires both frequency and coverage and so is naturally calculated at the zone-level. Like Shannon diversity, I also calculated native gamma diversity as a function of the total number of unique native species inventoried per zone as a whole, as opposed to the average number of unique native species per plot (referred to as native species richness) (Table 3, Appendix A).

While Shannon's diversity is commonly used in ecology (Guiden et al. 2021), previous American-based studies have established that C-value based metrics, specifically FQI, are among some of the most reliable metrics for predicting floral community quality overall (Rocchio 2007). Native FQI and AFQI are derived from a combination of species diversity counts and coefficients of conservatism, or "C-values" (Swink and Wilhelm 1979; 1994) Native FQI only considers native species, while AFQI is a single formula that includes introduced species as a term in the formula itself (Table 3; Rocchio 2007).

The C-value is a botanist-assigned value, mostly used in the USA, ascribed to every native plant species in a given region, and is specific to that region (Wilhelm and Rericha 2017). C-values offer a means of categorizing the relative "quality" or "desirability" of a species as part of a specific floral ecosystem, and a plant's resistance to environmental degradation and its ability to spread beyond its preferred habitat (Taft et al. 1997). When a given community is summarized in a species inventory, the average C-value of the community (also called "mean-C") presents a picture of the ecosystem's overall quality, with high C-values representing highly-specific species and relatively undamaged environments, and low C-values representing the opposite (Wilhelm and Rericha 2017). Remnant ecosystems typically have high mean-C communities, while disturbed ecosystems have low mean-C values (Wilhelm and Rericha 2017).

I categorized all species as either a native or introduced species (in the context of tallgrass mesic prairies) and assigned C-values ranging from 1 to 10 for native plants, and 0 for introduced species, according to Wilhelm and Rericha's *Flora of the Chicago Region* (Wilhelm and Rericha 2017; Bach and Kleiman 2021). All scientific names followed Mohlenbrock's *Vascular Flora of Illinois* (Mohlenbrock 2014; Appendix D). Assistance was also received from experienced land managers (Kobal 2025; Levernier 2025; Towey 2025). All unknown or unidentified species were excluded from metrics except species richness and Shannon's diversity index.

Table 3. The formulae used to calculate the per-plot and per-zone metrics used for statistical analysis. Once a value was calculated for all ten plots in a zone, they were averaged to get a value representative of the zone itself (excepting Shannon and gamma diversity, which require a zone-scale measurement).

Where i represents an individual from a given inventory (nat for native, all for all, and int for introduced species), Ci represents the C-value of an individual, $Covi.n$ is the number of instances of cover class of each individual species while n represents the cover class itself (classes range from 1 through 7, see Table 2; i.e., the sum of instances of cover class times their midpoint), $mp.n$ is the midpoint of each cover class, $total$ represents the total number of samples that a species could possibly appear in (i.e., 10 per zone), $frequi$ is the frequency of an individual species' appearance in replicate plots (i.e. how many times was this species sampled out of ten replicate plots), and $zone$ represents the summation across the entire zone, i.e. all ten plots. A sum of an individual term (e.g., $\sum Mcovi$) represents the summation of that term for all individual species. See Appendix A for full formulae.

* Denotes a metric that requires calculation but was not used for statistical analysis.

Parameter	Formula	Source
Species Richness (Native)	$N = \sum i(nat)$	Rocchio 2007
Species Richness (All)*	$S = \sum i(all)$	Rocchio 2007
Percentage Introduced Species	$Perc(int) = \left(\frac{\sum i(int)}{S} \right) * 100$	Rocchio 2007
Mean-C (Native)*	$MeanCn = \frac{\sum Ci}{N}$	Rocchio 2007
Native Floristic Quality Index (FQI)	$FQI = MeanCn * \sqrt{N}$	Taft, Hauser, and Robertson 2006
Adjusted Floristic Quality Index (AFQI)	$AFQI = \left(\frac{MeanCn}{10} * \frac{\sqrt{N}}{\sqrt{S}} \right) * 100$	Rocchio 2007
Mean Cover per Species*	$MCovi = \frac{\sum_{n=1}^7 ((\sum Covi.n) * mp.n)}{total}$	Rocchio 2007; Taft, Hauser, and Robertson 2006
Relative Coverage*	$RCovi = \frac{MCovi}{\sum MCovi}$	Taft, Hauser, and Robertson 2006
Relative Frequency*	$RFrequi = \frac{frequi}{\sum frequi}$	Taft, Hauser, and Robertson 2006
Relative Importance Value (IV 200)*	$p = \frac{RCovi + RFrequi}{\sum (RCovi + RFrequi)}$	Taft, Hauser, and Robertson 2006
Shannon-Wiener Diversity Index	$H' = - \sum p * \ln(p)$	Taft, Hauser, and Robertson 2006
Native Gamma Diversity	$\gamma = \sum i(nat).zone$	Sluis, Bowles, and Jones 2018

2.2 Secondary Data Sources

2.2.1 Land Cover Datasets and GIS Analysis

I used the United States Geological Survey's National Land Cover Database, or the "NLCD," as the basis of my land cover dataset (United States Geological Survey 2024). The NLCD is a pre-processed, 30 m resolution raster dataset that covers the entirety of the United States and classifies all land cover into sixteen categories (Table 4). I utilized the 2023 NLCD Annual dataset and applied standard GIS cleaning procedures, such as projecting (NAD 1983 (2011) UTM Zone 16N) and reclassifying. I simplified NLCD

classified groups from sixteen classes down to six (see Table 4, Appendix Figure B1). While zone outlines were hand drawn, site outlines were based upon data from the Prairie State Conservation Coalition's Illinois protected areas dataset (Holman and the Prairie State Conservation Coalition 2024).

Table 4. The breakdown of reclassified groups from the original NLCD categories (United States Geological Survey, 2024) to the simpler categories that I used for land cover analysis of rural and suburban landscape matrices.

* Open Field was ignored for the majority of analysis due to its common misidentification of land cover; see below for further explanation.

Number	Reclassified Category	Original NLCD Categories Included in Reclassification
1	Water	11: Open Water 12: Perennial Ice/Snow
2	Developed Land (Suburban)	21: Developed, Open Space 22: Developed, Low Intensity 23: Developed, Medium Intensity 24: Developed, High Intensity
3	Closed Ecosystems	41: Deciduous Forest 42: Evergreen Forest 43: Mixed Forest 52: Shrub/Scrub 90: Woody Wetlands
4	Open Ecosystems	31: Barren Land (Rock/Soil/Clay) 71: Grassland/Herbaceous 95: Emergent Herbaceous Wetlands
5	Farmland/Agricultural Usage (Rural)	82: Farmland
6*	Open Field (Ignored)	81: Pasture/Hay

I ignored classification 81, Pasture/Hay in the original NLCD dataset, as it regularly misidentified grassland (class 71) as farmland (class 82) or rangeland (class 81) and misclassified both remnant and restored prairie as pasture/hay. Furthermore, this classification had a particularly low NLCD-calculated confidence interval while NLCD classes 71 and 82 had significantly higher confidence intervals. Class 81 made up a relatively small percentage of total area, so I based my agricultural measurements only on NLCD class 82, pure farmland.

From the newly reclassified NLCD dataset, I calculated two types of land cover variables; binary land cover class and gradient land cover measures (Figure 7). Whether the rural or suburban class had a greater percentage of the area within a 2 km buffer around the *site* determined a site's binary land class. These binary land cover classes also serve as a statistical validation of a zone's original rural-suburban classification.

I assessed gradient land cover measurements by using Python and ArcGIS Pro (Esri 2024) to create a series of eight nested buffers (50 m, 200 m, 500 m, 1000 m, 1500 m, 2000 m, 2500 m, 3000 m; Figure 7) extending out from the boundaries of the *zone*. This is similar to methodology used in other grassland matrix studies (Davis, Debinski, and Danielson 2007).

Previous studies have demonstrated that plants respond to changes in the landscape matrix on a scale of around 2 km (Öckinger et al. 2012). However, some studies suggest that plants only meaningfully react to landscape matrix at larger landscape scales

(Dauber et al. 2003) while others suggest that a buffer of 500 m is enough to characterize landscape variables (Koper, Mozel, and Henderson 2010; Schwoertzig et al. 2016).

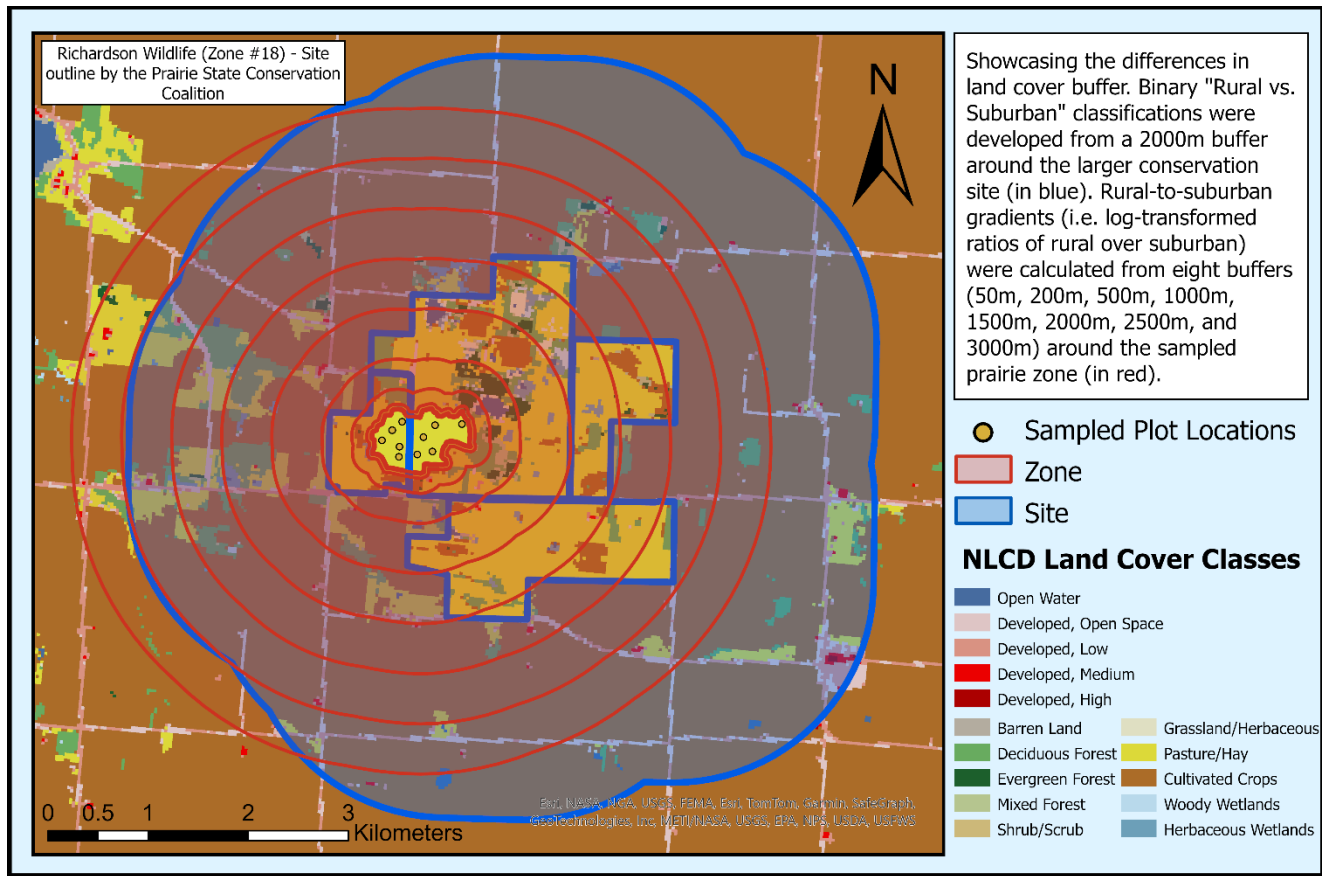


Figure 7. A map demonstrating the different methodology in measuring landscape matrix and land cover class surrounding a zone and site. Binary land cover classes were measured from buffers around the site, but more granular, gradient measures of land cover were drawn from buffers around the zone. While this map demonstrates the sixteen classes of the NLCD, land cover analyses were based upon six reclassified simple land cover classes, per Appendix Figure B1.

As there is some disagreement as to what scale is adequate to observe landscape matrix effects on vegetation, I deemed it important to consider different distances. From gradient nested buffers I was able to extract numerous additional landscape matrix variables, such as the exact area and percentage of each land cover class at each buffer distance. I could also take the average of all eight buffer distances for a more holistic, normalized approach to classifying surrounding matrix. I also measured the ratio of rural versus suburban cover class but since these data were right-skewed, I log-transformed all measured ratios before statistically analysis. This produced metrics that range between, on average, -7.5 and 2.5, with positive numbers representing areas that have a predominantly rural landscape matrix (i.e., rural land is greater than developed land), and negative numbers representing areas that have a predominantly suburban landscape matrix.

2.2.2 Management Survey Data

To understand how management might influence overall quality or interact with landscape matrix, I sent out email questionnaires to land managers directly responsible for every sampled zone. For a land manager to be applicable for the survey, they had to be directly responsible for the site (at least in part), have relevant educational experience in the field of grassland management, and express a desire to answer the survey.

These semi-structured questionnaire surveys (Arundel 2023; Thomander and Krosnick 2024) were designed to collect simple numeric values for a number of questions related to management variables associated with each site, including implemented management practices, intensity of management, staffing numbers, site and organizational operating budget, restoration and remnant status, age of restoration, special legal designations, and other qualities that make up part of a prairie's unique identity (Appendix C).

I received responses from 28 out of 30 zones (excluding only Midewin South Patrol Road and Midewin Glacial Plains). I defined age of restoration as the year of the survey (2024) minus the year given as the beginning of restoration (i.e., restoration begins 2006, age of restoration is 18). However, as true remnants have no "beginning," they were excluded from certain analyses of age. To account for this, sometimes responses were grouped into arbitrary categories for further analysis. Groups were designed to have largely even frequency. This also helped to fix issues of small sample size for certain questions.

This separation into distinct categories for statistical analysis was specifically applied to questions based on yes-no answers or questions with distinct multiple-choice answers (e.g., "What is the current estimated financial expenditure for only this site as a whole?"). Budgetary variables, age of restoration, and number of management methods were also categorized to see if distinct groups differed instead of looking for trends along a gradient.

For instance, remnant prairies have no true beginning and their age may range anywhere from 500 to 10,000 years, it was easier to compare them with grouping and use remnants as a distinct classification. I grouped remnant prairies as one classification and then grouped restored prairies into two classifications (younger than 30 years and between 30 and 100 years) to generate more even frequencies between groups.

The number of management methods was summed from respondent choices for methods they use, with options including burning, mowing, seeding, hand weeding, herbicide application, and an open-ended "other" option, for a maximum of six techniques. Management intensity was scored by assigning values to possible answers ranging from "Do not use this technique" (0) to "Several times a year" (5), and then summed to come up with a singular score representing management intensity. This score represents a summation of reported management frequency, intensity, and applied area for all reported management methods. A similar technique was used for the summation of staffing presence.

Within the questions of the survey, site budget asks respondents to estimate the amount of money directly spent on a single site (with a preference for funds spent specifically on management), while operating organization budget asks participants to estimate what the total annual operating budget is for the organization that owns the site itself. This includes all funds, including those beyond just prairie management.

Participants were informed about their rights to data use and anonymity prior to filling out the survey, and their responses will remain anonymous.

2.3 Statistical Analysis

Using scripts in R, RStudio (RStudio Team 2020), and ggplot2 (Wickham 2016), I compared six response variable floral metrics (native species richness, percent introduced species, native FQI, AFQI, Shannon's diversity, and gamma diversity) and sixty-eight landscape and management predictor variables to explore potential trends and correlations. I dropped one outlier zone (zone 3, Trelease Prairie) as it was a mixed-use research and conservation site, as well as the southernmost site. Its values were not comparable to other zones (Appendix Figure B16).

I utilized the standard p-value limit of 0.05 to test for significance, and $0.05 < p < 0.10$ for near-significance. I tested normality with the Shapiro-Wilk test. Categorical-Continuous comparison tests varied based on number of categories and normality; two-category normal comparisons used T tests, while non-normal comparisons used Wilcoxon rank sums test. Normal comparisons with greater than two categories used ANOVA testing, and non-normal comparisons utilized the Kruskal-Wallis test. Comparisons between two continuous datasets used an ANOVA linear model for normal comparisons, and a Spearman's correlation test was performed on non-normal comparisons. However, an ANOVA linear model was still used to get R^2 fit values for non-normal comparisons (referred to as $R^2(\text{LM})$ in this text). I discovered that mixed modelling was not applicable to this dataset; see Appendix Figure B18 for a full explanation.

Besides the log transformation of rural-suburban ratio, I also log-transformed a number of other ratios of landscape matrix variables, such as ratios of wilderness to suburban or open wilderness to closed wilderness. Beyond this, no other values were transformed after being measured. See Appendix A, Table A3 for a full list of analysed variables.

3 Results

3.1 Anthropogenic Landscape Matrix

For all response variables except percent introduced, both floral diversity and floral quality increased as plots became more rural; the inverse is true as sites become more suburban (Figure 8; Figure 9). This trend was still present even in cases where results were not statistically significant. There were no correlations at all between the response variables and the amount of surrounding wilderness, either open (grassland) or closed (forest), in the landscape matrix.

That being said, in order to observe significant trends, it was typically not enough to classify zones as either rural or suburban. For example, only native species richness ($t_{(24,31)}=2.0903$, $p=0.047$) demonstrated a significant response when classified by simple binary rural-suburban groupings (Figure 8). Shannon diversity was near-significant ($W=146$, $p=0.077$).

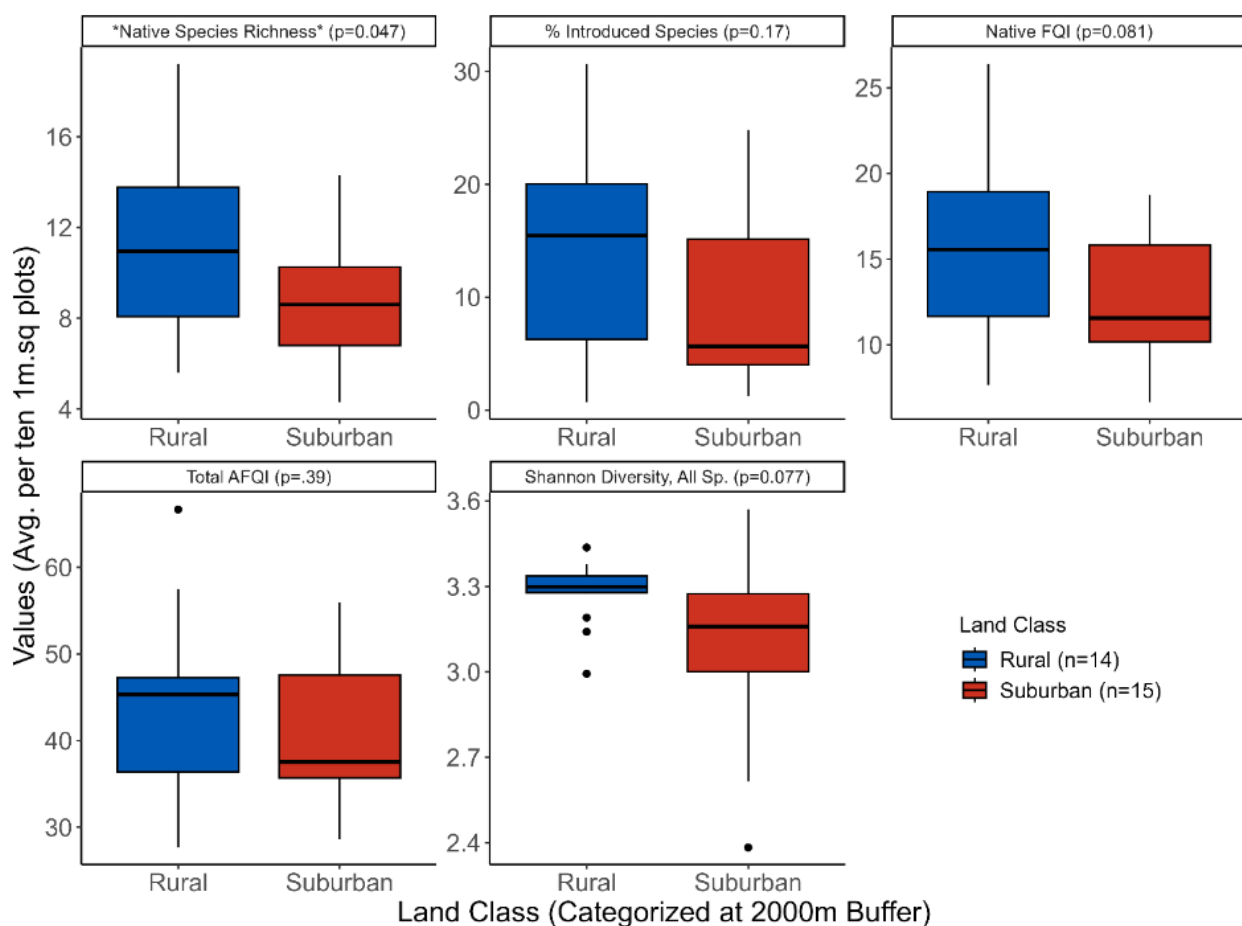


Figure 8. Boxplots depicting the five response variables against a binary classification of surrounding landscape matrix. “Suburban” represents a site where, at a 2000 m buffer, the percentage of developed landscape was greater than agricultural landscape. “Rural” represents the inverse of this, a space where agricultural surroundings outclassed developed space. Asterisks around a variable name denote statistical significance.

However, when comparing response variables across a gradient of rural-to-suburban landscape matrix, trends became more consistent than the responses from binary classification. From four of the five response variables I observed positive quality trends

(Figure 9) associated with an increasing percentage of gradient average rural percentage. Native species richness ($F_{(n: 1, d: 27)}=9.85$, $p=0.0041$, $R^2=0.27$), FQI ($F_{(n: 1, d: 27)}=10.27$, $p=0.0035$, $R^2=0.28$) and AFQI ($F_{(n: 1, d: 27)}=4.68$, $p=0.040$, $R^2=0.15$.) all saw gains in quality and diversity as areas became more rural. Shannon diversity was also significant ($\rho_{(27)}=0.43$, $p=0.019$, $R^2(\text{LM})=0.12$). Percent of introduced species was not significant.

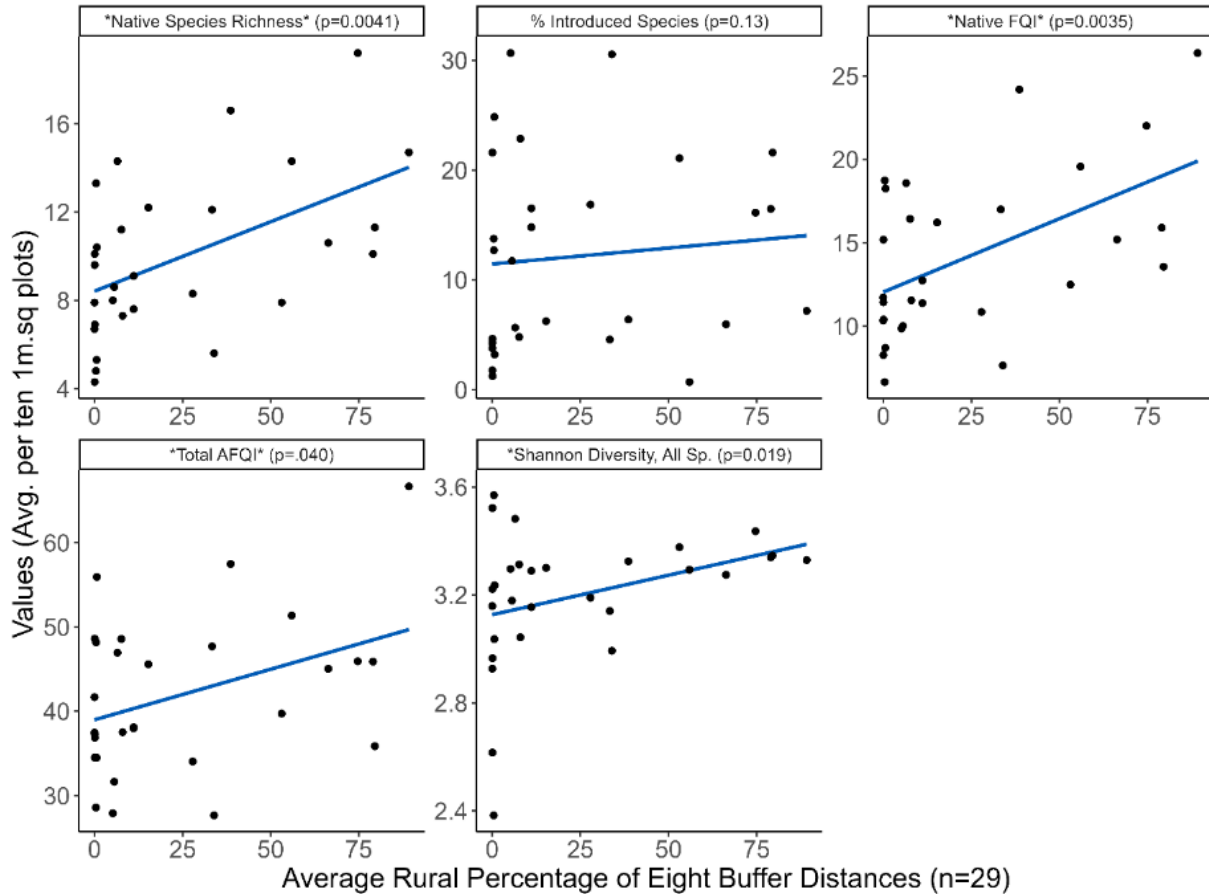


Figure 9. Line graphs depicting the five response variables against the average percentage of rural landscape, as drawn from eight different buffer distances around the zone (50m, 200m, 500m, 1000m, 1500m, 2000m, 2500m, and 3000m). This gradient approach to landscape matrix follows a buffer directly from the zone itself, *not* the site, meaning that it takes into consideration more complex matrices.

The inverse of this gradient also had similar responses, in that quality generally decreased as the percentage of suburban landscape matrix increased. While native species richness ($F_{(n: 1, d: 27)}=6.34$, $p=0.018$, $R^2=0.19$), FQI ($F_{(n: 1, d: 27)}=4.56$, $p=0.042$, $R^2=0.14$), and Shannon diversity ($\rho_{(27)}=-0.43$, $p=0.020$, $R^2(\text{LM})=0.14$) maintained significance, AFQI was no longer significant when comparing suburban percentage.

Comparing the average log-transformed ratio of rural versus suburban coverage against response variables demonstrated increases in quality consistent with the established trend. This was true for native species richness ($F_{(n: 1, d: 25)}=7.63$, $p=0.011$, $R^2=0.23$), native FQI ($F_{(n: 1, d: 25)}=6.15$, $p=0.020$, $R^2=0.20$), and Shannon diversity ($\rho_{(25)}=0.42$, $p=0.031$, $R^2(\text{LM})=0.16$), but not AFQI or percent introduced species (Figure 10). There were no differences in trend between the average ratio and individual ratios at different buffer distances (Appendix Figure B2).

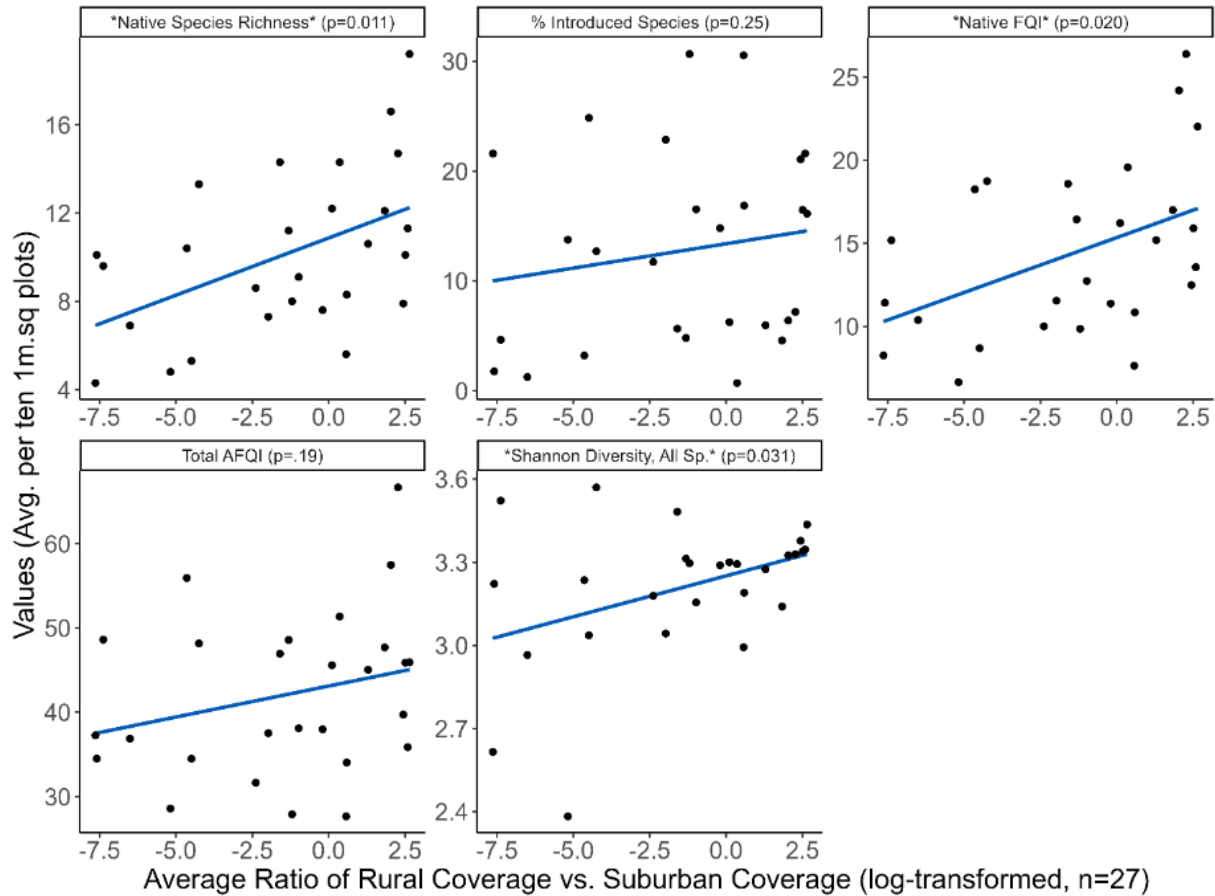


Figure 10. Line graphs depicting the five response variables against the log-transformed average ratio of rural to suburban landscape, as drawn from eight different buffer distances around the zone (50m, 200m, 500m, 1000m, 1500m, 2000m, 2500m, and 3000m). These values range from positive values, representing a greater ratio of rural land, to negative values, representing a greater ratio of suburban land.

Notably, the percent of introduced species increased with increased rural landscape matrix (Figures 9, 10, and 11). This inverse is also true and significant for changes in suburban landscape matrix; where other quality metrics decrease as percent of suburban matrix increases, the percent of also introduced species decreases (i.e., a quality increase) ($\rho_{(27)}=-0.41$, $p=0.028$, $R^2(\text{LM})=0.14$) (Figure 11). Other response variables decrease in quality as the average of suburban matrix increases.

Of important note for these results and others is that these are just trends; while p values may be statistically significant, R^2 values are consistently poor for all analysed variables in both landscape matrix and in management/socioeconomic variables. Variance in the data is very large, and results can be interpreted as general trends but cannot be used as strong predictors of quality or numeric estimates. See section 4 for further interpretation of this.

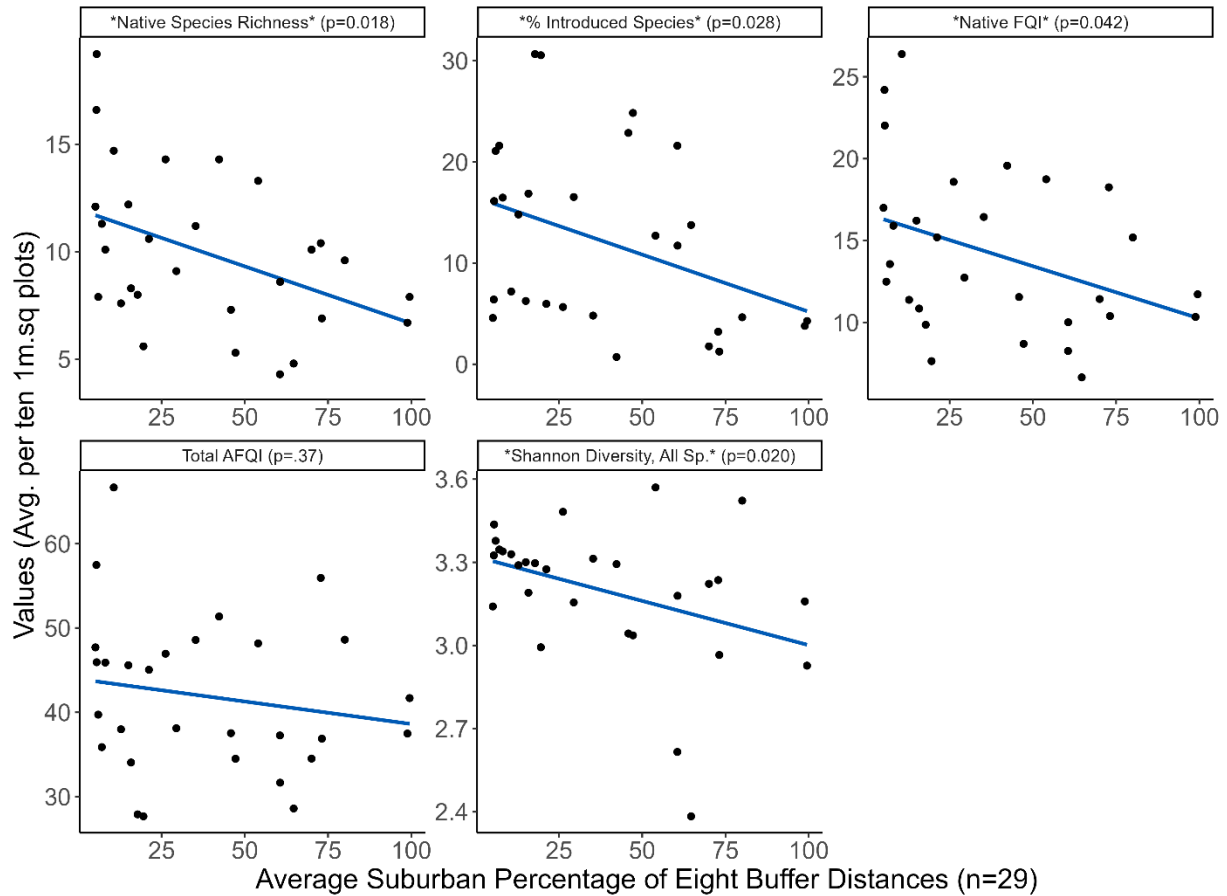


Figure 11. Line graphs depicting the five response variables against the average percentage of suburban matrix across eight different buffer distances around the zone (50m, 200m, 500m, 1000m, 1500m, 2000m, 2500m, and 3000m). This supports the observed trend that increasing rural matrix increases quality, as it demonstrates the inverse that increasing suburban matrix decreases quality. The notable exception is percent introduced species, which demonstrates a quality increase (decreased introduced species) with increasing suburban matrix.

3.2 Management

Introduced species decreased as the number of management interventions increased ($\chi^2_{(2)}=9.94$, $p=0.0069$), and there was a general but non-significant increase in quality and diversity of other metrics as the number of utilized management methods increased. Zones that used three or fewer management methods typically saw a higher percentage of introduced species, though there was no statistical difference between zones that used four methods and zones that used five or six methods (Figure 12).

When breaking management methods down into the presence or intensity of individual management techniques, only the presence and intensity of hand weeding was correlated to a noticeable increase in quality, regardless of the presence/absence of other

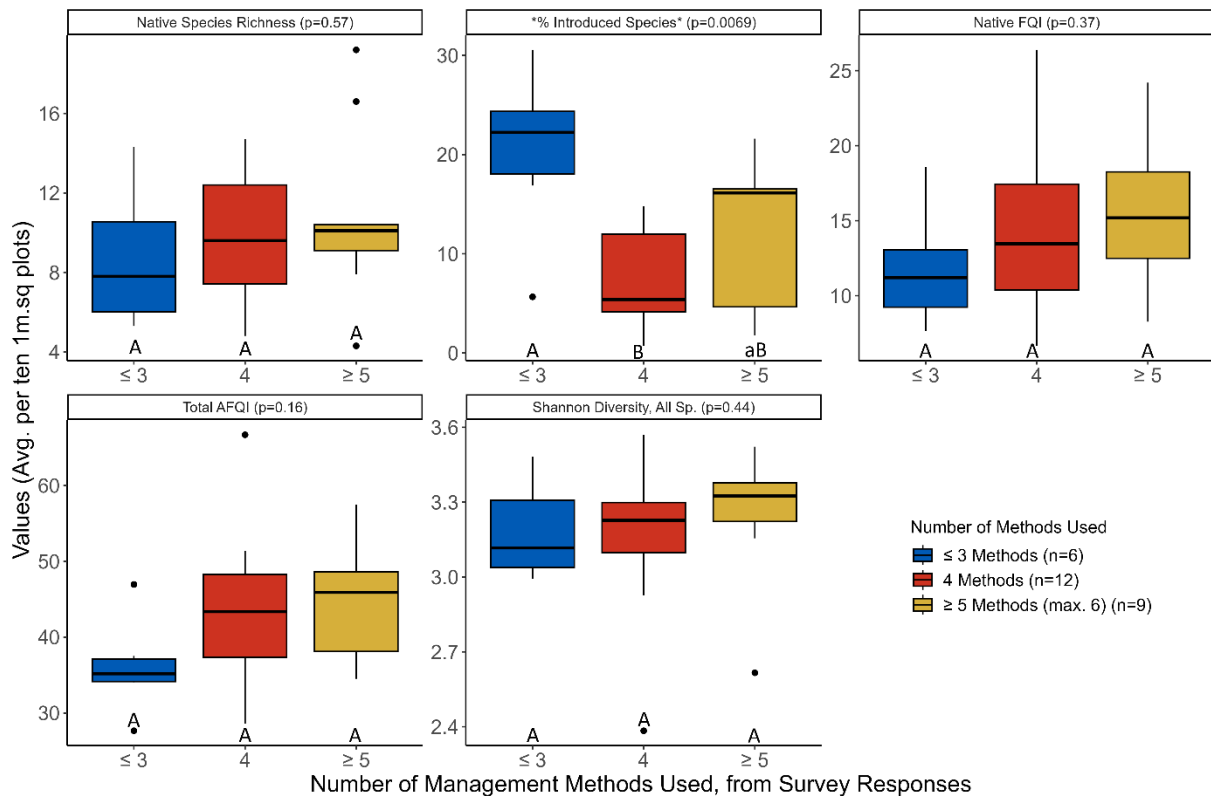


Figure 12. A series of bar graphs demonstrating the five response variables against the number of management methods used. The number of management methods have been grouped based on relative frequency to give groups that are roughly evenly-sized, allowing for more consistent sample sizes.

management methods. This increase from hand weeding was true for all five response variables: native species richness ($t_{(22,23)}=-2.75$, $p=0.012$), percent of introduced species ($W=120$, $p=0.046$), native FQI ($t_{(23,47)}=-3.22$, $p=0.0037$), AFQI ($t_{(22,66)}=-3.51$, $p=0.0019$), and Shannon diversity ($W=36$, $p=0.020$), with all seeing increases in quality and diversity with the presence of hand weeding (Figure 13). No other individual management technique displayed this trend (Appendix Figure B10), though herbicide and fire could not be tested for presence/absence as nearly every site used some iteration of these two techniques (26 out of 27 sites for both herbicide and burning). Intensity of techniques similarly only saw significant responses to hand weeding.

The number of staff, i.e., of human labor available to a site/zone, showed a positive relationship with native FQI ($F_{(n: 1, d: 22)}=4.12$, $p=0.055$, $R^2=0.16$) and AFQI ($F_{(n: 1, d: 22)}=4.41$, $p=0.047$, $R^2=0.17$) (Appendix Figure B3). However, this was only true for the staffing score as a composite measure of permanent staff, seasonal staff, volunteers, and hired contractors; there was no notable response when comparing individual components, such as just permanent staff or just contractor presence/absence.

The total sum of management saw no consistent trends or correlations with the five primary response metrics (Appendix Figure B13). Quality and diversity generally increased as the intensity and application of management increased, but results were not significant and the spread of response variables was generally very wide. Only percent of introduced species was near significance, showing a decrease in introduced species as management sum increased ($\rho_{(25)}=-0.33$, $p=0.097$, $R^2(\text{LM})=0.19$).

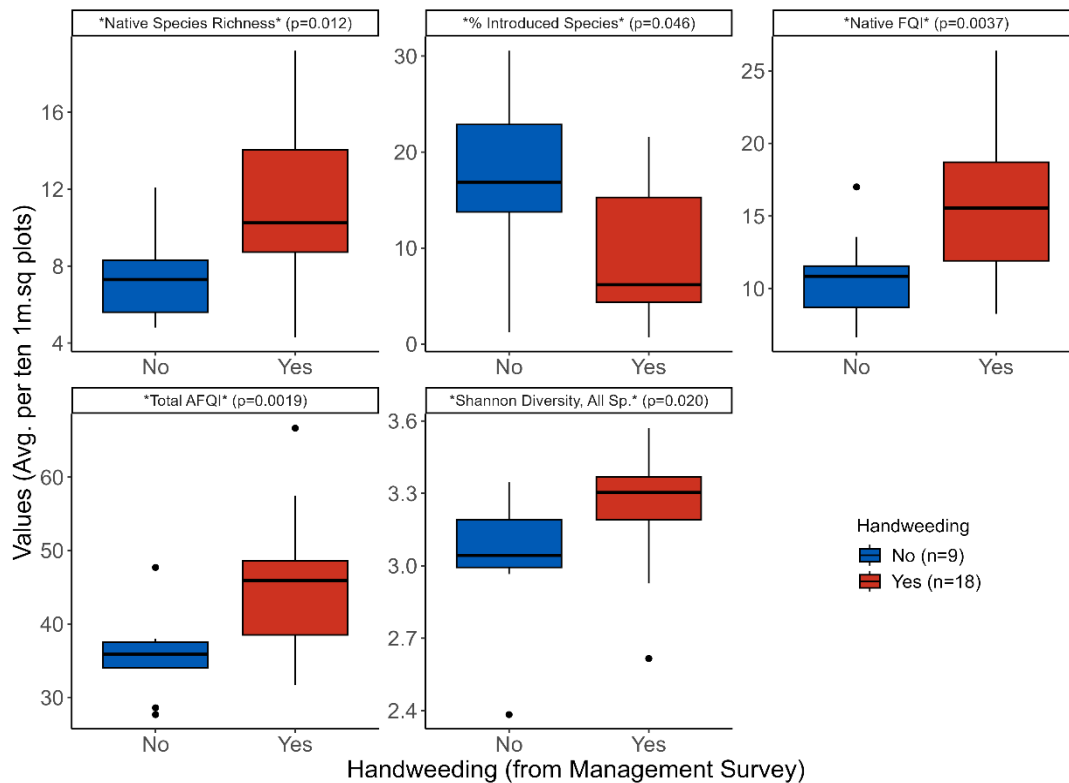


Figure 13. A series of bar graphs demonstrating the five response variables against the presence or absence of hand weeding as a management method, taken independent of the presence or absence of other management methods. No other management methods displayed this kind of positive trend in quality by their presence or absence alone.

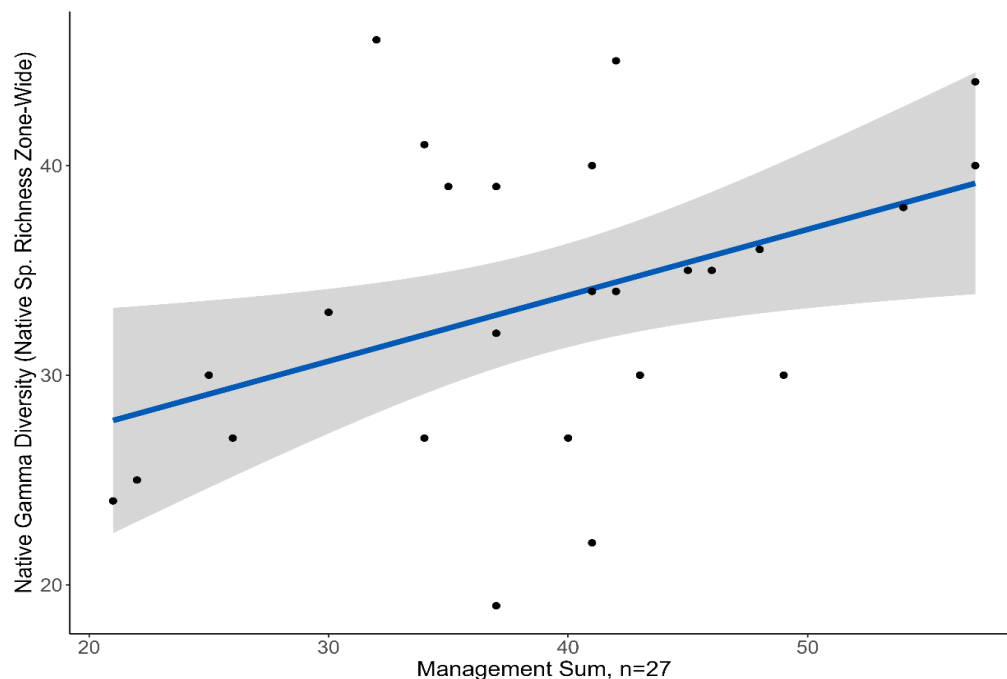


Figure 14. A line graphs demonstrating the gamma diversity (the total number of unique species inventoried per zone) against sum of management as measured from the management survey. This takes into account the frequency, intensity, and area of coverage for all management methods to give a general sense of the management applied to a given site/zone. Gamma diversity was the only response variable to demonstrate a significant trend to management sum (Appendix Figure B13).

Interestingly, native gamma diversity saw a significant positive correlation with management (Figure 14, $F_{(n: 1, d: 25)}=5.51$, $p=0.027$, $R^2=0.18$). Native gamma diversity also shared similar positive quality trends as a response to hand weeding sum ($F_{(n: 1, d: 25)}=7.82$, $p=0.0098$, $R^2=0.24$) and staffing sum ($F_{(n: 1, d: 22)}=5.90$, $p=0.024$, $R^2=0.21$) but otherwise did not respond to any other predictor variables.

3.3 Age, Budget, and Other Socioeconomic Factors

3.3.1 Age of Restoration

The percent of introduced species was lower in prairies that had a longer restoration history (Figure 15, $\rho_{(20)}=-.44$, $p=0.041$, $R^2(\text{LM})=0.17$), though this analysis only included restoration, not remnants. When ages were grouped in order to test for differences between remnant prairies and restored prairies, there were no significant differences between differently-aged prairies (Appendix Figure B4). There was also no significant difference between groups of fully remnant, fully restoration, or mixed prairies.

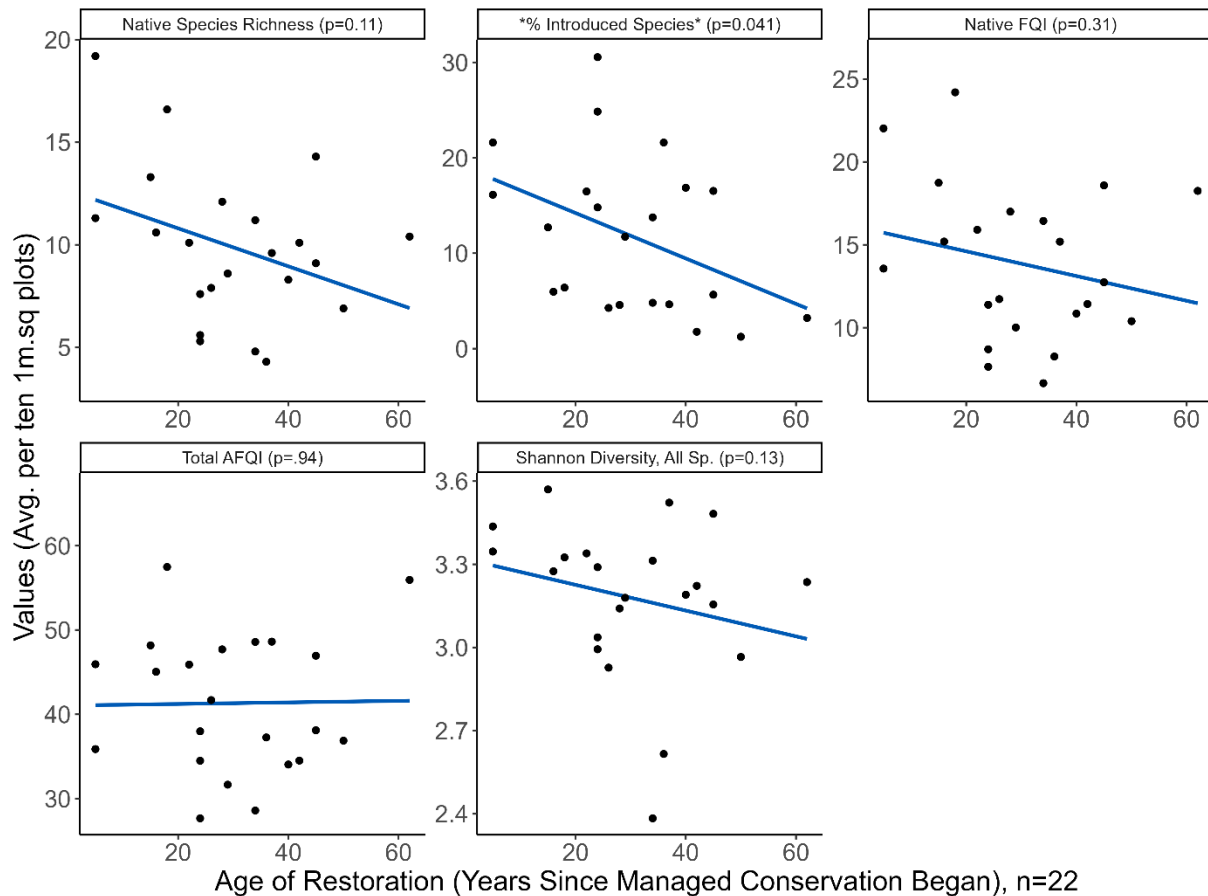


Figure 15. Line graphs demonstrating the five response variables when plotted against the known ages of restoration. This is effectively measuring the responses to the number of years since active conservation and restoration began. As remnant prairies have existed prior to written history, they have been dropped from this particular comparison. However, comparisons where zones were grouped by age (with remnants included) yielded no significant trends.

3.3.2 Organizational Funding

When the estimated budget per site was normalized against hectare of zone (i.e., *site* budget divided by *zone* size), response variables generated positive trends in quality as site budget per zone hectare increased. This was only significant for AFQI ($F_{(n: 1, d: 23)}=4.47$, $p=0.046$, $R^2=0.16$), and near significance for native FQI ($F_{(n: 1, d: 23)}=3.89$, $p=0.061$, $R^2=0.14$), though this trend was generally present but non-significant for all five response variables (Appendix Figure B5). There were no correlations when comparing *site* budget divided by *site* size or when grouping site budgets without accounting for size.

In contrast, smaller total annual operating budgets of overhead organizations (the group responsible for a given zone; see Appendix E) were correlated to higher prairie quality (Figure 16). This was true for native species richness ($F_{(2)}=8.09$, $p=0.0023$), native FQI ($F_{(2)}=8.62$, $p=0.0017$), AFQI ($F_{(2)}=5.23$, $p=0.014$), and Shannon diversity index ($\chi^2_{(2)}=6.95$, $p=0.031$), but not percent of introduced species. Only low-budget groups were significantly different.

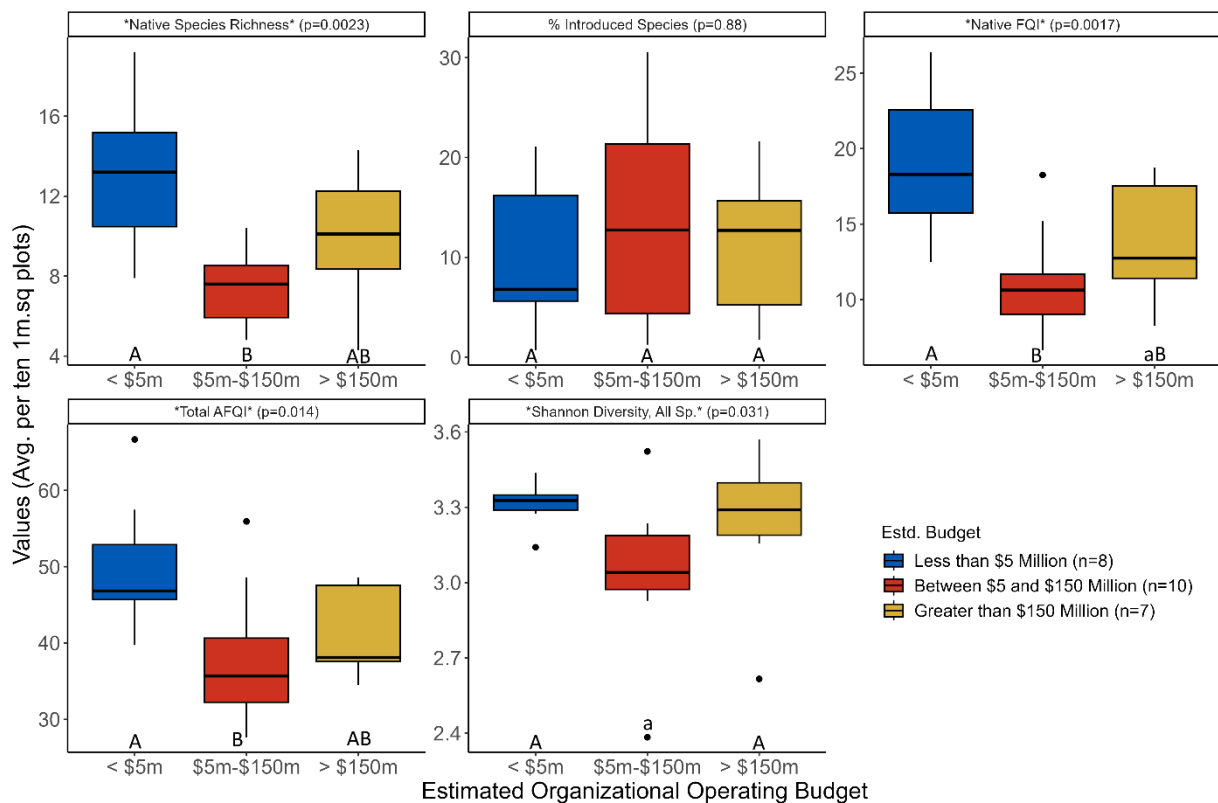


Figure 16. Bar graphs demonstrating the five response variables when plotted against the survey-derived overhead organizational operating budget. This represents the estimated annual budget of the organization directly responsible for the management and personnel of the zone in question. There is a distinct difference between organizations with relative low budgets versus and organizations with mid-range budgets, though this difference is less pronounced when compared to high-budget organizations.

3.3.3 Ownership and Legal Designation

There were no statistical differences between zones with special legal protections and zones without; both had similar quality and diversity (Appendix Figure B12). Conversely, whether or not a zone was under private or public ownership demonstrated significant differences, with quality generally being higher in zones that were under private ownership. This was true for native species richness ($t_{(10.59)}=2.55$, $p=0.028$), percent of introduced species ($W=22$, $p=0.027$), native FQI ($t_{(5.98)}=2.90$, $p=0.027$), and AFQI ($t_{(5.40)}=3.37$, $p=0.018$), but not Shannon diversity index (Figure 17).

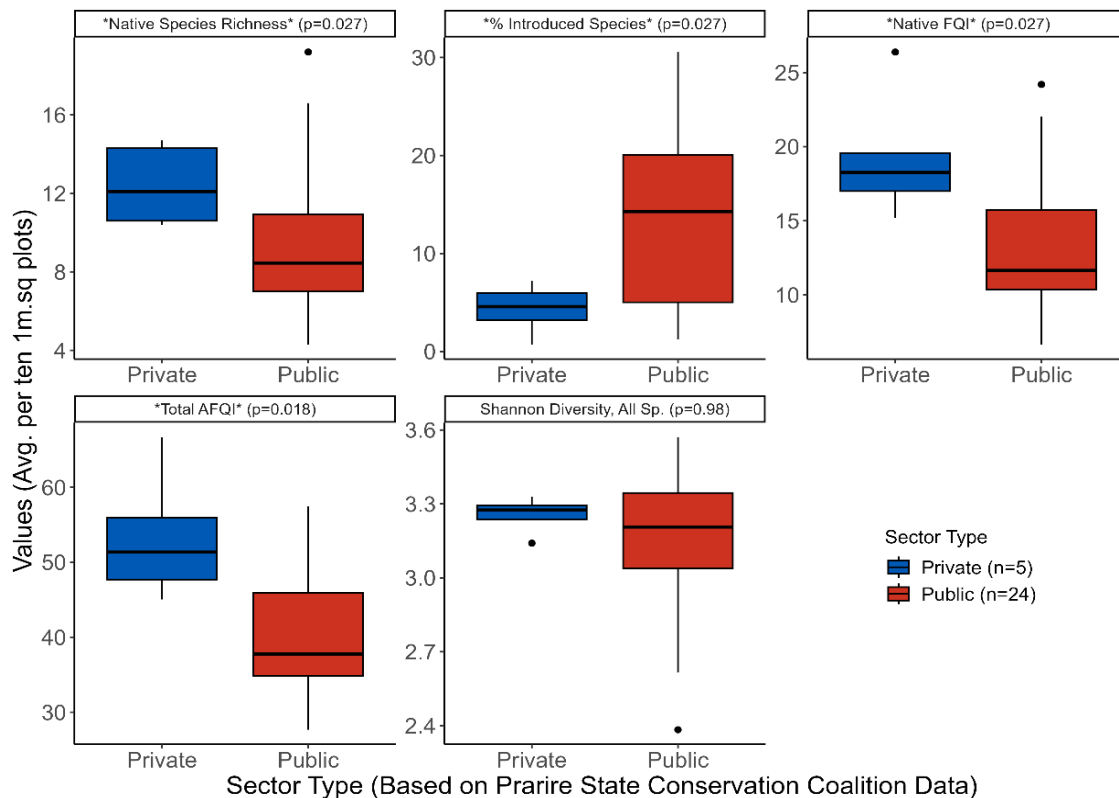


Figure 17. Bar graphs demonstrating the five response variables when plotted against land ownership classification, as determined by the Prairie State Conservation Coalition. In general, quality was higher in zones that were under private ownership (i.e., by a dedicated conservation organization) than under public ownership (such as municipal, state, or federal ownership). However, the sample size of privately-owned sites/zones is low when compared to publicly-owned sites/zones.

3.3.4 Area

The only trend observed in relation to the area of both the zone and the area of the site overall is that the percent of introduced species tended to increase as the size of the zone increased ($\rho_{(27)}=0.45$, $p=0.016$, $R^2(\text{LM})=0.060$). Beyond this, non-significant trends suggest a decrease in quality as zones increase in size.

3.4 Multivariate Modelling and Sub-Grouping

Running very large multivariate generalized linear models suggested that the three most impactful explanatory predictors for variance in native FQI ($n=23$) were, in order: the site budget per hectare of zone, the log-transformed ratio of rural to suburban, and the survey-derived sum of hand weeding intensity. After trimming down the model to only

contain those three most relevant variables, this trend was still true. (Appendix Table A2).

From a model comparing the five primary response variables (and Gamma diversity) and those three predictor variables, I determined that rural-suburban ratio was a significant explanatory factor for all response variables except AFQI and gamma diversity. Hand weeding sum was a significant explanatory variable for percent introduced species and gamma diversity. Site budget per hectare of zone was significant for native species richness and native FQI (Appendix Table A2).

Classifying response variables based on sub-groups (e.g., presence of hand weeding, special legal designations) produces results consistent with previously-established trends and no new correlations of interest (Appendix Figure B6, B11, B12). Breaking response variables into smaller sub-groups generally resulted in a loss of significance, unless those subgroups were based on landscape matrix (Appendix Figure B6, B7, B8).

Grouping response variables based on landscape matrix and then comparing them against a landscape matrix predictor demonstrated that response values are more similar among rural areas than they among suburban areas. Rural areas are more likely to see a statistically significant trend than are suburban areas when grouped (Appendix Figure B7). For example, when examining the ratio of rural-suburban landscape while responses are grouped by landscape type, rural-grouped Shannon diversity ($\rho_{(12)}=0.67$, $p=0.011$, $R^2(\text{LM})=0.14$) had a strong trend, and rural-grouped species richness ($F_{(n: 1, d: 12)}=3.45$, $p=0.089$, $R^2=0.22$), rural-grouped FQI ($F_{(n: 1, d: 12)}=4.34$, $p=0.059$, $R^2=0.27$), and rural-grouped AFQI ($F_{(n: 1, d: 12)}=3.53$, $p=0.085$, $R^2=0.23$) are near significance. No suburban groupings detected any significance.

4 Discussion

I would summarize my primary results in the following way: in general, results correlated increasing rural matrix with increasing quality. Increased rural landscape matrix also correlated to increased percentage of introduced species. Quality and diversity generally increased as more management methods were used, but the inclusion of hand weeding as a management technique had the strongest correlation to increased quality. A greater individual site budget per hectare of zone correlated to increased quality. Multivariate analysis suggested that the variables with the greatest correlation to prairie quality are the log-transformed ratio of rural to suburban, the individual site budget per hectare of the zone, and the survey-derived sum of hand weeding intensity.

From this, I interpret three primary conclusions. Firstly, the quality of prairies is generally higher in rural areas than it is in suburban areas. Secondly, while an increased number of applied management techniques generally improves the quality of a space, external socioeconomic factors associated with the application of management can dramatically complicate the picture. Inputs of human labor may instead be a better indicator of higher quality prairies rather than measurements correlated to number of management techniques or their application intensity. Thirdly, generalized linear models suggest that both landscape and management factors affect prairie quality in an unequal but comparable magnitude. While there is relatively little research on comparisons between types of anthropogenic landscape matrix or external socioeconomic factors, these results do generally match expectations based on literature examining other components of prairie health and management.

That being said, while many correlations and trends have statistically significant P-values, not a single measured regression result has a R^2 value greater than 0.3. The highest reported R^2 values are for the responses to the average percentage of rural area, which hovers around 0.27. The data reflects correlations and trends, but making specific numeric claims would be difficult; perhaps the variations are too large, or the sample size may be too small. Nevertheless, my results are consistent with established research. Broad conclusions can still be made from the data primarily due to the consistent, reliable responses generated across several iterations of testing, especially in anthropogenic landscape matrices. While certain broad limitations should be considered, I analysed several distinct floral metrics and found consistent results in a relatively novel field.

Furthermore, although surveys may be a variable and sometimes bias-prone form of data collection (Arundel 2023; Thomander and Krosnick 2024), my approach to this research was interdisciplinary and combined several forms of data collection, including species inventories and geographic analysis as well as land manager surveys. This kind of interdisciplinary approach can be very powerful, and is not often carried out in tallgrass prairie research. I believe that the consistency of my analyses and conclusions lends credence to this research approach.

4.1 Anthropogenic Landscape Matrix

My results suggest that increasingly rural anthropogenic landscapes were more conducive to higher floral diversity and quality. Suburban or developed landscapes, on the other hand, were correlated with decreasing floral quality and diversity. This is in agreement with other studies that examined conserved grassland responses to landscape matrix and ecosystems along urbanization gradients (Bascompte and Rodríguez 2001; Williams et al. 2005; Rowe 2013; Cochard et al. 2019). These results were consistent in both binary and more comprehensive gradient measures of landscape matrix.

One could extrapolate this to mean that rural prairies are more stable. The increase in native diversity is correlated to higher resilience against disturbance (Allison 2008; Brudvig 2011; Rice et al. 2018). In addition, this increase of floral quality, per coefficients of conservatism and FQI, also represents an increased specificity of species composition (Swink and Wilhelm 1994; Wilhelm and Rericha 2017). In other words, rural prairies are more likely to contain more diverse and specialized prairie plant assemblages than suburban prairies. They are therefore more likely to play stronger ecological roles, likely increasing faunal diversity and ecosystem services output as well (Steinauer and Collins 1996; McAlpine et al. 2016; Bach and Kleiman 2021; Guiden et al. 2021).

Unexpectedly, I found that the amount of surrounding wilderness seemed to have no impact on prairie quality. This is surprising because previous studies have established that surrounding prairie cores with wilderness matrix tends to maintain or increase quality (Robertson, Anderson, and Schwartz 1997; Rowe 2013; Kemmerling et al. 2022). However, my observed response is consistent with the results of Grman, Bassett, and Brudvig (2013) who claimed that landscape matrix only rarely influenced the quality and outcome of a prairie restoration. However, they only considered agricultural and wilderness landscape types and did not measure sites within suburban landscape matrices, as I did in my study. Perhaps, based on their findings and mine, the most important landscape distinction for tallgrass prairies is between different types of anthropogenic landscapes, and not between natural and anthropogenic landscapes.

In my study, I found that response variables in rural prairies were both slightly more consistent and possessed a greater likelihood for higher quality when compared to response measurements in suburban prairies. I suspect this difference is due to the fact that threats to conserved spaces are largely consistent across the homogenous Illinois rural monoculture landscape (typically either corn or soy), while suburban Illinois landscapes are far more heterogenous in their matrix and the threats present therein (Robertson, Anderson, and Schwartz 1997; Ricketts and Imhoff 2003; Patz et al. 2004; Gámez-Virúes et al. 2015; Seidl and Golobič 2020; Priyadarshana et al. 2024). Some of the largest threats from rural landscapes are generally land conversion, introduced species invasion, pesticide drift, agricultural runoff, and fragmentation (Steinauer and Collins 1996; Robertson, Anderson, and Schwartz 1997; Patz et al. 2004; Koper, Mozel, and Henderson 2010). While Illinois suburban landscapes may, in theory, offer increased diversity from their increased heterogeneity, these suburban landscapes face most of these same pressures that rural landscapes do as well as other threats, such as increased industrial pollution, much higher direct human disturbance, and much stronger barriers to connectivity, such as larger paved areas and decreased open soil (Ricketts and Imhoff 2003; Patz et al. 2004; Wilby and Perry 2006; Palacio et al. 2018; Theodorou 2022).

Suburban landscapes likely create a more difficult landscape matrix than rural landscapes, potentially influencing the trends observed in this research. My results support this conclusion, suggesting it less likely for suburban prairies to meet the same quality levels as rural prairies.

All this does not mean that suburban prairies are worth less than rural prairies; some studies demonstrate that conservation in urban areas can be even more important than conservation in rural areas (Ricketts and Imhoff 2003). I hope that my results may plant a potential seed of research into the kinds of floral changes that occur along gradients of anthropogenic landscape matrices. Tallgrass prairies are exceptionally heterogenous landscapes, adapted to numerous landscape-scale disturbance regimes and with hundreds of native plant species (Steinauer and Collins 1996; Robertson, Anderson, and Schwartz 1997). This was confirmed by my findings as I identified over 190 unique native plant species between my 30 sampled prairie zones, some of which I only identified in one or two zones. I propose that, despite difficulties with anthropogenic landscape matrix, conservationists and city planners should jump at any available opportunities to

protect or restore prairie, regardless of location. We have, and can continue to develop, a complex, heterogeneous patchwork of prairies across a wide variety of landscape matrices. If well managed, all prairies will continue to provide a strong base for maintaining biodiversity across a wide area, and continue studying changes between prairies in different anthropogenic landscapes.

4.2 Management and Socioeconomic Considerations

The most surprising of these results might be the correlations related to hand weeding. Previous studies have typically not measured the total intensity of management, but rather the impacts of one or two singular types of management technique, usually grazing or fire (Steinauer and Collins 1996; Wagle and Gowda 2018). My analyses looked at a larger set of management techniques, and I found that neither the presence/absence or the intensity of fire or mowing (the closest grazing substitute according to Römermann et al. 2009) had any correlation to plant species diversity or quality. Instead, I observed that the presence of hand weeding as a management technique very strongly did, which is consistent with previous studies but perhaps surprising in its comparative strength (Clark and Wilson 2001; Kimball et al. 2013; Ede and Greet 2021).

I would suggest that it is not hand weeding itself that is responsible for the increase (though it certainly could be), but that increased human labor leads to increased quality. Sites that have greater labor input are probably more likely to use hand weeding as part of their management and are also more likely to spend more total labor hours on prairie management in general (Smith et al. 2010). In my study, this was confirmed by an increased staff size (i.e., human labor) being correlated to increased quality and diversity. One could say that the input of human labor is more strongly correlated to prairie quality than is the exact intensity of management.

This could also link to my study's lack of correlation with age or remnant status. This trend does go against established literature (Steinauer and Collins 1996; Allison 2008; Grman, Bassett, and Brudvig 2013; Rowe 2013; Bach and Kleiman 2021). However, I would a possible solution: ultimately, it could be because, first, older prairie restorations may reach a theoretical cap on species establishment, and second, restorations today typically apply higher diversity seed mixes and maybe having higher quality management overall when compared to older restorations (Grman, Bassett, and Brudvig 2013; Campbell and Hooymans 2016; Bach and Kleiman 2021; Chakravorty, Harrington, and Bach 2024). Newer restorations may see higher-quality management and labor, and thus have higher floral quality in return.

The other most surprising trend I found is that low-budget organizations seemed to have higher quality overall than sites run by organizations with very large operating budgets. One possible explanation for this is that organizations with smaller budgets may be incentivized to increase the cost-effectiveness of their restoration and thus, in turn, increase the actual outcome of their restoration (Ribaud et al. 2017). Allocation of limited funds, especially for the mitigation of land-use conflicts, can be one of the major limiting factors for conservation success (Remme and Schröter 2016; Alagador and Cerdeira 2018). Small-budget organizations with singular conservation priorities may be more able to focus on the greatest return on investment or smartest allocation of resources than very large, disparate organizations with many distinct priorities and stakeholders (Donlan, Luque, and Wilcox 2015; Shoo et al. 2021). This is especially possible when previous literature has established that government conservation management programs may not always be the best at allocating resources where they are most needed (Ribaud et al. 2017).

Similarly, in their survey responses, some managers suggested that, despite an organization's large budget, few resources were actually allocated to land management and conservation. Large budgets may mask the fact comparatively little financial resources are invested in restoration or management. This would explain the strong trend towards site budget per hectare of zone being correlated with increased quality; smaller sites with comparatively higher conservation-allocated budgets will have higher quality, even if the operating organization's total budget is small. Further research into the exact breakdown of budget by site and management may yield new conclusions, but suffice it to say that budget does not determine quality.

One final surprising result was that sites with legal designations (i.e., suggesting higher quality and therefore necessitating greater protection) were indistinguishable from sites without legal designations. In Illinois, legal designations are granted to sites that are high quality, home to endangered species, or represent some sort of unique feature that necessitates permanent protection (Natural Areas Evaluation Committee 2023; Illinois Department of Natural Resources 2025). The fact that legally-designated high-quality natural spaces are indistinguishable from non-designated spaces suggests that there are numerous high-quality sites that are legally unprotected. The number of sites with legal protections may need to be broadened to include all sites of equivalently-high quality. As most legally-protected sites are managed by government entities, these external variables are not always directly correlated to success.

From all of this, I would conclude that measures of human labor first, and measures of management second, are the best guarantee of a prairie's quality and diversity. Conversely, budget and operating organization are not a guarantee. Unsurprisingly, management, and to an extent human labor, is still the most-correlated socioeconomic variable to a prairie's success. Successful prairies need focused plans with effective resource allocation, but most of all, successful prairies need experienced professionals. If large organizations ignore the need for experts who can manage focused, applied labor, then conserved prairies will suffer regardless how large the operating organization is.

4.3 Multivariate Analyses

I suggest that both landscape matrix, management, and socioeconomic considerations all serve to explain the variance of sites, but they explain the variance in unequal quantities and for different measures of quality and diversity. It is very likely that there are countless intertwined variables which affect prairie floral quality in numerous ways, of which these variables are only some (Steinauer and Collins 1996; Grman, Bassett, and Brudvig 2013; Wagle and Gowda 2018). My model suggests that management and external socioeconomic variables are likely a greater determining factor of overall quality and diversity than landscape matrix. As much as landscape matrix may serve as an influential factor in the quality of a site, there is more than can be done through management to affect prairie health. If the surrounding landscape matrix becomes more urbanized, then increased management may be needed to compensate for the detrimental effects of increasing urbanization. Awareness of floral quality changes related to external variables can help land managers plan management in their prairies.

4.4 Future Directions

One future consideration for anthropogenic landscape matrices is further classification of sub-types of developed environments, as well as defining the types of and relative intensity of their threats. There is a wide variety of matrix subtypes within the suburban landscape that are uncaptured in my study, such as green residential areas or industrial

zones. Defining a gradient of low-intensity development versus high intensity development, or specific types of surrounding development, could prove to be fruitful research. Future studies may also benefit from assigning more variables at the plot, not zone, level, and using more complex statistical models, such as nonmetric multidimensional scaling, permutation ANOVA, variation partitioning, and structural equation modelling (Grman, Bassett, and Brudvig 2013; Koerner et al. 2016; Back and Kleiman 2021; Guiden et al. 2021; Chakravorty, Harrington, and Bach 2024). Changes in species composition between rural and suburban prairies could also enhance our understanding of the changes in prairie communities in different anthropogenic landscape matrices.

A more detailed study of socioeconomic variables and land ownership practices could also expand upon the conclusions drawn here, and could have important social applications and not just ecological uses. For example, many studies suggest that returning ecosystems to indigenous management, or at least engaging indigenous peoples more directly in management, improves both floral and faunal diversity as well as the wellbeing of modern indigenous cultures (Gardner et al. 2022; Beattie et al. 2023; Parks and Tsioumani, 2023). However, there are often no consistent ways to implement changes in management or ownership. This is especially true in the United States, where land ownership and natural resources management is typically of a patchwork nature between a wide array of organizations. Illinois is currently actively returning some state-owned natural spaces to indigenous management (O'Connor 2025). If this trend is to continue, further explorations into the impacts of socioeconomic differences between land management variables and ownership styles can help ensure the highest quality conservation systems within available resource constraints.

4.5 Synthesis and Conclusions

There is little research on both the differences between tallgrass prairies in different types of anthropogenic landscape matrices and the potential impacts of the larger socioeconomic variables beyond straight land management. Through this study, I have helped to fill a knowledge by determining that there is a measurable difference in quality and diversity between prairies in different anthropogenic landscape matrices. My research also suggests that increasing inputs of human labor is correlated to increasing the quality and diversity of conserved prairies. This research highlights new trends in relatively understudied parts of prairie management. They also highlight deficiencies in how prairies are managed in Illinois today.

Synthesizing these conclusions into one general idea, I would say that all prairies are unique, have unique needs, and could contribute to the complexity of the remaining tallgrass prairie landscape. However, what tallgrass prairies need to succeed today is increased individual focus. Between remnants, restorations, rural prairies, suburban prairies, privately-owned sites, publicly-owned sites, varying levels of legal protections, and a vast array of other background factors, no two prairies are the same. The best, most successful prairies are the ones that get increased labor and attention in order to succeed in spite of their difficult surroundings.

In a very broad sense, all prairies and managers should play to their site's unique strengths and weaknesses. Suburban prairies, despite their difficulties, have a vastly larger pool of potential human labor to draw from. They are in a much better position to draw on that value of human labor than rural prairies. Managers for these prairies may benefit from looking outward, to volunteers, to seasonal workers, to getting local people excited about prairies. Excited people may be more willing to support government plans to increase prairie management and staffing, too. And rural prairies, despite not having the people, have the space. If the edges of prairies are the problem, then perhaps a large, connected preserve that takes advantage of that ample space is the best direction for rural

prairies to go in. If the landscape matrix is gentler around rural prairies, then it may be best to take advantage of it.

Urban spaces are only growing, and farming is becoming ever more resource-intensive, and prairies across the state are going to need proper management now more than ever. This does not necessarily require a rethinking of prairie management at a level of land managers, but perhaps a rethinking of prairie management at the level of landowners, operating organizations, and government. As I heard again and again from my surveys and interactions with land managers, managers today are often overworked and undersupplied. If we want to see these places improve, not just languish, we need to support the people who keep these prairies healthy, the land managers themselves. Rural vistas, big budgets, and legal protections do always not guarantee a healthy or diverse prairie, despite what lawmakers might believe otherwise. There is no singular solution to prairie management, but the people who put in the actual labor are the most qualified to make decisions regarding prairie management. As a society, we need to recognize that what prairies need the most are people.

Tallgrass prairies are enchantingly heterogenous landscapes, but they have vanished almost entirely; they are one of the most endangered ecosystems on Earth. We cannot afford to lose more of them to the detrimental effects of a difficult landscape matrix. More labor may be necessary, but the rewards are worth it in both rural and suburban spaces.

If we want to save the prairies as we know them today, then we need to treat them as they are now. Each Illinois prairie is an individual plot with unique requirements, and they are as much a part of their surroundings as their surroundings are a part of them. While increased labor is maybe the best bet for better restoration, what this means for each individual prairie will always differ. With enough labor, care, and experience, prairie management can provide what is uniquely needed for each prairie. And, in turn, each prairie can provide something unique not just to floral diversity and quality, but to our lives and the landscapes in which we live.

Acknowledgements and Attributions

I want to acknowledge the myriad indigenous peoples and cultures who originally lived, continue to live, and will live, on these prairies. This includes, but is not limited to, the Kaskaskia, Kiikaapoi, Meskwaki, Myaami, Peoria, Potawatomi, Sauk, and Ochethi Sakowin (Sioux) peoples. I would like to thank every land manager who answered my survey request, as well as everyone else who pointed me in the right direction of who to ask. While I can't thank you all personally here (because of privacy requirements), your contributions were invaluable to the success of my thesis. Thank you for the work you do for the prairies. I would, more specifically, like to thank the land managers and organizations who worked closely with me during the surveying period: Elizabeth Bach and the management team at Nachusa Grasslands; Jamie Ellis and the University of Illinois at Urbana-Champaign's Natural Areas program, Walter Levernier and the management team at Fermi National Accelerator Laboratory; Hector Ortiz Cano and the Chicago Botanic Garden's Negaunee Institute for Plant Conservation Science and Action; and J. Brian Towey and the Richardson Wildlife Foundation. Thank you for your guidance, time, consideration, and the use of your space. I would also like to thank Scott Kobal and the Forest Preserve District of DuPage County for assistance in species classification. Furthermore, my deepest gratitude to my advisor, Dr. Anna Treydte, whose direction and assistance made this project possible. Thanks also to Heather Wood, for statistics consultation. Finally, I'd like to thank my family, classmates, friends, and my partner Sandra, all of whose consistent support, both emotional and educational, have been crucial to this process. Thank you.

I declare that there are no conflicts of interest associated with this research or the conclusions drawn therein. Generative AI was not used to create any part of this report or its appendices.

References

- Abrams, M.D. and Hulbert, L.C., 1987. Effect of topographic position and fire on species composition in tallgrass prairie in northeast Kansas. *American Midland Naturalist*, pp.442-445.
- Adelman, C. and Schwartz, B., 2013. *Prairie Directory of North America: The United States, Canada, and Mexico*. Oxford University Press.
- Alagador, D. and Cerdeira, J.O., 2018. A quantitative analysis on the effects of critical factors limiting the effectiveness of species conservation in future time. *Ecology and Evolution*, 8(6), pp.3457-3467.
- Allen, M.S. and Palmer, M.W., 2011. Fire history of a prairie/forest boundary: more than 250 years of frequent fire in a North American tallgrass prairie. *Journal of Vegetation Science*, 22(3), pp.436-444.
- Allison, S.K., 2008. Differences in the effects of drought upon restored and remnant prairies (Illinois). *Ecological Restoration*, 26(2), pp.95-97.
- Ampleman, M.D., Crawford, K.M. and Fike, D.A., 2014. Differential soil organic carbon storage at forb-and grass-dominated plant communities, 33 years after tallgrass prairie restoration. *Plant and soil*, 374, pp.899-913.
- Anderson, R.C., 2006. Evolution and origin of the Central Grassland of North America: climate, fire, and mammalian grazers. *The Journal of the Torrey Botanical Society*, 133(4), pp.626-647.
- Arundel, A., 2023. How to design, implement, and analyse a survey. In *How to Design, Implement, and Analyse a Survey*. Edward Elgar Publishing.
- Avenza Systems Inc., 2024 *Avenza Maps: Offline Mapping (5.3.3) [Mobile App]*. Available at: <https://apps.apple.com/us/app/avenza-maps-offline-mapping/id388424049> (Accessed: August and September 2024)
- Avolio, M.L., Komatsu, K.J., Collins, S.L., Grman, E., Koerner, S.E., Tredennick, A.T., Wilcox, K.R., Baer, S., Boughton, E.H., Britton, A.J. and Foster, B., 2021. Determinants of community compositional change are equally affected by global change. *Ecology Letters*, 24(9), pp.1892-1904.
- Bach, E.M. and Kleiman, B.P., 2021. Twenty years of tallgrass prairie restoration in northern Illinois, USA. *Ecological Solutions and Evidence*, 2(4).
- Baer, S.G., Blair, J.M., Collins, S.L. and Knapp, A.K., 2003. Soil resources regulate productivity and diversity in newly established tallgrass prairie. *Ecology*, 84(3), pp.724-735.
- Barak, R.S., Williams, E.W., Hipp, A.L., Bowles, M.L., Carr, G.M., Sherman, R. and Larkin, D.J., 2017. Restored tallgrass prairies have reduced phylogenetic diversity compared with remnants. *Journal of Applied Ecology*, 54(4), pp.1080-1090.
- Bascompte, J. and Rodríguez, M.Á., 2001. Habitat patchiness and plant species richness. *Ecology Letters*, 4(5), pp.417-420.
- Batt, B.D.J., 1996. "Prairie Ecology – Prairie Wetlands", in Samson, F.B. and Knopf, F.L (ed.) *Prairie Conservation: Preserving North American's Most Endangered Ecosystem*. Washington, D.C.; Island Press, pp.77-88.
- Beattie, M., Fa, J.E., Leiper, I., Fernández-Llamazares, Á., Zander, K.K. and Garnett, S.T., 2023. Even after armed conflict, the environmental quality of Indigenous Peoples' lands in biodiversity hotspots surpasses that of non-Indigenous lands. *Biological Conservation*, 286, p.110288.
- Bragg, T.B. and Steuter, A.A., 1996 "Prairie Ecology – The Mixed Prairie", in Samson, F.B. and Knopf, F.L (ed.) *Prairie Conservation: Preserving North American's Most Endangered Ecosystem*. Washington, D.C.; Island Press, pp.53-65.
- Brudvig, L.A., 2011. The restoration of biodiversity: where has research been and where does it need to go?. *American journal of botany*, 98(3), pp.549-558.
- Campbell, R.E. and Hooymans, J.L., 2016. Results from four decades of successional prairie restoration and an update on ecological land management at Fermilab in Batavia, Illinois (No. FERMILAB-CONF-16-167-FESS). Fermi National Accelerator Lab.(FNAL), Batavia, IL (United States).
- Chadde, S. W., 2019. *Prairie Plants of Illinois*. United States of America: Orchard Innovations.

- Chakravorty, J., Harrington, J.A. and Bach, E.M., 2024. Bison Grazing in Eastern Tallgrass Prairie Does Not Alter Plant Diversity after Five Years. *Natural Areas Journal*, 44(4), pp.215-222.
- Chalcraft, D.R., Williams, J.W., Smith, M.D. and Willig, M.R., 2004. Scale dependence in the species-richness–productivity relationship: the role of species turnover. *Ecology*, 85(10), pp.2701-2708.
- Chapman, K.A. and Brewer, R., 2008. Prairie and savanna in southern Lower Michigan: history, classification, ecology. Michigan Botanical Club.
- Chicago Metropolitan Agency for Planning, 2024. City of Chicago's Boundaries, ArcGIS Online Map Viewer. [Dataset] Available at: <https://www.arcgis.com/home/item.html?id=312a5a22ff7d4c56875d41fbb7e6f0c3&sublayer=0>. (Accessed 3 Aug 2024)
- Clark, C.M., 2011. Great plains, chapter 11. In: Pardo, L.H. and Robin-Abbott, M.J., 2011. Assessment of nitrogen deposition effects and empirical critical loads of nitrogen for ecoregions of the United States. Gen. Tech. Rep. NRS-80. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station. 291 p., 80, pp.1-291.
- Clark, D.L. and Wilson, M.V., 2001. Fire, mowing, and hand-removal of woody species in restoring a native wetland prairie in the Willamette Valley of Oregon. *Wetlands*, 21(1), pp.135-144.
- Cochard, A., Pithon, J., Braud, F., Beaujouan, V., Bulot, A. and Daniel, H., 2019. Intraspecific trait variation in grassland plant communities along urban-rural gradients. *Urban Ecosystems*, 22, pp.583-591.
- Coleman, A.L. and Wait, D.A., 2022. Urban Prairie Plots and Gardens Can Sustain Plant-Pollinator Interactions Similar to Established Rural Prairies. *The American Midland Naturalist*, 188(1), pp.102-118.
- Collins, S.L., 2000. Disturbance frequency and community stability in native tallgrass prairie. *The American Naturalist*, 155(3), pp.311-325.
- Corbett, E.A., 2004. A comparison of Illinois remnant prairies, 1976 to 1988.
- Coulloudon, B., Eshelman, K., Gianola, J., Habich, N., Hughes, L., Johnson, C., Pellant, M.L., Podborny, P., Rasmussen, A., Robles, B. and Shaver, P.L., 1999. Sampling vegetation attributes: interagency technical reference.
- Craine, J.M., Nippert, J.B., Towne, E.G., Tucker, S., Kembel, S.W., Skibbe, A. and McLaughlan, K.K., 2011. Functional consequences of climate change-induced plant species loss in a tallgrass prairie. *Oecologia*, 165, pp.1109-1117.
- Dauber, J., Hirsch, M., Simmering, D., Waldhardt, R., Otte, A. and Wolters, V., 2003. Landscape structure as an indicator of biodiversity: matrix effects on species richness. *Agriculture, Ecosystems & Environment*, 98(1-3), pp.321-329. <https://www.sciencedirect.com/science/article/pii/S0167880903000926>
- Davis, J.D., Debinski, D.M. and Danielson, B.J., 2007. Local and landscape effects on the butterfly community in fragmented Midwest USA prairie habitats. *Landscape Ecology*, 22, pp.1341-1354. <https://link.springer.com/article/10.1007/s10980-007-9111-9>
- Donlan, C.J., Luque, G.M. and Wilcox, C., 2015. Maximizing return on investment for island restoration and species conservation. *Conservation Letters*, 8(3), pp.171-179.
- Ede, F. and Greet, J., 2021. Post-sowing weed control technique can affect woody seedling numbers, with early hand-weeding potentially more beneficial than early spraying. *Ecological Management & Restoration*, 22(3), pp.266-273.
- Esparrago, J. and Kricsfalussy, V., 2015. Traditional grassland management and surrounding land use drive the abundance of a prairie plant species in urban areas. *Landscape and Urban Planning*, 142, pp.1-6.
- Esri, 2024. ArcGIS Pro (3.2.2) [Software]. Redlands, CA: Environmental Systems Research Institute.
- Fiaschi, T., Fanfarillo, E., Maccherini, S., Bacaro, G., Bonari, G., Foggi, B., Peruzzi, L., Pinzani, L., Rosati, L., Scoppola, A. and Viciani, D., 2023. Effectiveness of different metrics of floristic quality assessment: The simpler, the better?. *Ecological Indicators*, 149, p.110151.
- Fowler, A., Basso, B., Maureira, F., Millar, N., Ulbrich, R. and Brinton, W.F., 2024. Spatial patterns of historical crop yields reveal soil health attributes in US Midwest fields. *Scientific Reports*, 14(1), p.465.
- Gámez-Virués, S., Perović, D.J., Gossner, M.M., Börschig, C., Blüthgen, N., De Jong, H., Simons, N.K., Klein, A.M., Krauss, J., Maier, G. and Scherber, C., 2015. Landscape

- simplification filters species traits and drives biotic homogenization. *Nature communications*, 6(1), p.8568.
- Gardner, E.M., Puad, A.S.A., Pereira, J.T., anak Tagi, J., anak Nyegang, S., Miun, P., Jumian, J., Pokorny, L. and Zerega, N.J., 2022. Engagement with indigenous people preserves local knowledge and biodiversity alike. *Current Biology*, 32(11), pp.R511-R512.
- Google LLC., 2022. Google Earth Pro (7.3.6) [Computer Program] Available at: <https://www.google.com/earth/about/versions/>
- Grman, E., Bassett, T. and Brudvig, L.A., 2013. EDITOR'S CHOICE: Confronting contingency in restoration: management and site history determine outcomes of assembling prairies, but site characteristics and landscape context have little effect. *Journal of Applied Ecology*, 50(5), pp.1234-1243.
- Guiden, P.W., Barber, N.A., Blackburn, R., Farrell, A., Fliginger, J., Hosler, S.C., King, R.B., Nelson, M., Rowland, E.G., Savage, K. and Vanek, J.P., 2021. Effects of management outweigh effects of plant diversity on restored animal communities in tallgrass prairies. *Proceedings of the National Academy of Sciences*, 118(5), p.e2015421118.
- Harmon-Threatt, A. and Chin, K., 2016. Common methods for tallgrass prairie restoration and their potential effects on bee diversity. *Natural areas journal*, 36(4), pp.400-411.
- Higgins, K.F., Oldemeyer, J.L., Jenkins, K.J., Clambey, G.K. and Harlow, R.F., 1996. Vegetation sampling and measurement. *Research and management techniques for wildlife and habitats*, 5, pp.567-591.
- Holman, D., and the Prairie State Conservation Coalition, 2024. I-View: Protected Natural Lands & Municipal Parks. [Web Map] Available at: <https://pscc.maps.arcgis.com/apps/webappviewer/index.html?id=fb8dbaacc7b04c7fa12e850b9e7141ab> (Accessed 3 Aug 2024)
- Howe, H.F., 1994. Managing species diversity in tallgrass prairie: assumptions and implications. *Conservation Biology*, 8(3), pp.691-704.
- Illinois Department of Agriculture, 2025. Facts about Illinois Agriculture. [Webpage] Available at: <https://agr.illinois.gov/about/facts-about-illinois-agriculture.html> (Accessed May 2025)
- Illinois Department of Natural Resources, 2024. Nature Preserves Directory. [Webpage] Available at: <https://dnr.illinois.gov/inpc/directory/naturepreservedirectory.html> (Accessed 3 Aug 2024)
- Illinois Department of Natural Resources, 2025. Illinois Nature Preserves Commission: Protection. [Webpage]. Available at: <https://dnr.illinois.gov/inpc/protection.html> (Accessed May 2025)
- iNaturalist, 2024. iNaturalist Classic (3.3.4) [Mobile App]. Available at: <https://apps.apple.com/us/app/inaturalist-classic/id421397028> (Accessed: August and September 2024)
- iNaturalist, 2024. Seek by iNaturalist (2.16.1) [Mobile App]. Available at: <https://apps.apple.com/gb/app/seek-by-inaturalist/id1353224144> (Accessed: August and September 2024)
- Jules, E.S. and Shahani, P., 2003. A broader ecological context to habitat fragmentation: why matrix habitat is more important than we thought. *Journal of Vegetation Science*, 14(3), pp.459-464.
- Kaul, A.D. and Wilsey, B.J., 2021. Exotic species drive patterns of plant species diversity in 93 restored tallgrass prairies. *Ecological Applications*, 31(2), p.e2252.
- Kemmerling, L.R., Rutkoski, C.E., Evans, S.E., Helms IV, J.A., Cordova-Ortiz, E.S., Smith, J.D., Vázquez Custodio, J.A., Vizza, C. and Haddad, N.M., 2022. Prairie strips and lower land use intensity increase biodiversity and ecosystem services. *Frontiers in Ecology and Evolution*, 10, p.833170.
- Kim, J.S. and Noto, A.E., 2025. Intraspecific diversity of multiple plant species shows no change across an urbanization gradient. *Urban Ecosystems*, 28(1), pp.1-9.
- Kimball, S., Lulow, M.E., Mooney, K.A. and Sorenson, Q.M., 2014. Establishment and management of native functional groups in restoration. *Restoration Ecology*, 22(1), pp.81-88.
- Knapp, A.K. and Smith, M.D., 2001. Variation among biomes in temporal dynamics of aboveground primary production. *Science*, 291(5503), pp.481-484.
- Knapp, A.K., Blair, J.M., Briggs, J.M., Collins, S.L., Hartnett, D.C., Johnson, L.C. and Towne, E.G., 1999. The keystone role of bison in North American tallgrass prairie: Bison increase habitat heterogeneity and alter a broad array of plant, community, and ecosystem processes. *BioScience*, 49(1), pp.39-50.

- Knapp, A.K., Burns, C.E., Fynn, R.W., Kirkman, K.P., Morris, C.D. and Smith, M.D., 2006. Convergence and contingency in production–precipitation relationships in North American and South African C 4 grasslands. *Oecologia*, 149, pp.456-464.
- Kobal, S., 2025. Email to Scott Kobal, 5 February.
- Koerner, S.E., Avolio, M.L., La Pierre, K.J., Wilcox, K.R., Smith, M.D. and Collins, S.L., 2016. Nutrient additions cause divergence of tallgrass prairie plant communities resulting in loss of ecosystem stability. *Journal of Ecology*, 104(5), pp.1478-1487.
- Koper, N., Mozel, K.E. and Henderson, D.C., 2010. Recent declines in northern tall-grass prairies and effects of patch structure on community persistence. *Biological Conservation*, 143(1), pp.220-229. <https://www.sciencedirect.com/science/article/pii/S0006320709004431>
- Kurz, D., 2021. *Illinois Wildflowers*. Jasper, Arkansas: Tim Ernst Publishing.
- Landi, S. and Chiarucci, A., 2010. Is floristic quality assessment reliable in human-managed ecosystems?. *Systematics and Biodiversity*, 8(2), pp.269-280.
- Larson, D.L., Hernández, D.L., Larson, J.L., Leone, J.B. and Pennarola, N., 2020. Management of remnant tallgrass prairie by grazing or fire: Effects on plant communities and soil properties. *Ecosphere*, 11(8), p.e03213.
- Levernier, W. G., 2025. Email to Walter Levernier, 12 February.
- Mackelprang, R., Grube, A.M., Lamendella, R., Jesus, E.D.C., Copeland, A., Liang, C., Jackson, R.D., Rice, C.W., Kapucija, S., Parsa, B. and Tringe, S.G., 2018. Microbial community structure and functional potential in cultivated and native tallgrass prairie soils of the Midwestern United States. *Frontiers in microbiology*, 9, p.1775.
- McAlpine, C., Catterall, C.P., Nally, R.M., Lindenmayer, D., Reid, J.L., Holl, K.D., Bennett, A.F., Runting, R.K., Wilson, K., Hobbs, R.J. and Seabrook, L., 2016. Integrating plant-and animal-based perspectives for more effective restoration of biodiversity. *Frontiers in Ecology and the Environment*, 14(1), pp.37-45.
- Mohlenbrock, R. H., 2014. *Vascular Flora of Illinois: A Field Guide*, 4th ed. Carbondale, Illinois: Southern Illinois University Press.
- Moilanen, A. and Arponen, A., 2011. Administrative regions in conservation: balancing local priorities with regional to global preferences in spatial planning. *Biological Conservation*, 144(5), pp.1719-1725.
- Natural Areas Evaluation Committee, 2023. The Standards and Guidelines for the Illinois Natural Areas Inventory [Digital Document]. Available at: https://naturalheritage.illinois.gov/content/dam/soi/en/web/naturalheritage/documents/INAI%20Standards%20and%20Guidelines_NAEC%20Approved%202-28-23.pdf (Accessed May 2025)
- O'Connor, J., 2025. "Illinois returns stolen land to Prairie Band Potawatomi Nation," AP News, Mar. 25, Available at: <https://apnews.com/article/potawatomi-illinois-treaty-reservation-805f5030d9abbac93ce3da3cec18a636>
- Öckinger, E., Lindborg, R., Sjödin, N.E. and Bommarco, R., 2012. Landscape matrix modifies richness of plants and insects in grassland fragments. *Ecography*, 35(3), pp.259-267. <https://nsojournals.onlinelibrary.wiley.com/doi/full/10.1111/j.1600-0587.2011.06870.x>
- Palacio, F.X., Ibañez, L.M., Maragliano, R.E. and Montalti, D., 2018. Urbanization as a driver of taxonomic, functional, and phylogenetic diversity losses in bird communities. *Canadian Journal of Zoology*, 96(10), pp.1114-1121.
- Parks, L. and Tsoumani, E., 2023. Transforming biodiversity governance? Indigenous peoples' contributions to the Convention on Biological Diversity. *Biological Conservation*, 280, pp.109933-1.
- Patz, J.A., Daszak, P., Tabor, G.M., Aguirre, A.A., Pearl, M., Epstein, J., Wolfe, N.D., Kilpatrick, A.M., Foutopoulos, J., Molyneux, D. and Bradley, D.J., 2004. Unhealthy landscapes: policy recommendations on land use change and infectious disease emergence. *Environmental health perspectives*, 112(10), pp.1092-1098.
- Petipas, R.H., Peru, C., Parks, J.M., Friesen, M.L. and Jack, C.N., 2024. Prairie soil improves wheat establishment and accelerates the developmental transition to flowering compared to agricultural soils. *Canadian Journal of Microbiology*, 70(11), pp.482-491.
- Polley, H.W., Derner, J.D. and Wilsey, B.J., 2005. Patterns of plant species diversity in remnant and restored tallgrass prairies. *Restoration Ecology*, 13(3), pp.480-487.
- Priyadarshana, T.S., Martin, E.A., Sirami, C., Woodcock, B.A., Goodale, E., Martínez-Núñez, C., Lee, M.B., Pagani-Núñez, E., Raderschall, C.A., Brotons, L. and Rege, A., 2024. Crop and landscape heterogeneity increase biodiversity in agricultural landscapes: A global review and meta-analysis. *Ecology Letters*, 27(3), p.e14412.

- QGIS Association, 2024. QGIS Geographic Information System (3.36.2) [Computer Program]
Available at: <http://www.qgis.org/>
- Reemts, C. and Belcher, B., 2022. Making a Good Prairie Better: Plant Diversity Increased in a Formerly Hayed Remnant Prairie Now Managed with Fire. *Natural Areas Journal*, 42(1), pp.18-27.
- Reinhold, Rich, 2025. "New Standards and Geographic Definitions for Metropolitan Statistical Areas." Illinois Department of Employment Security. Available at: <https://ides.illinois.gov/resources/labor-market-information/geographic-definitions.html> (Accessed May 2025).
- Remme, R.P. and Schröter, M., 2016. Effects of budget constraints on conservation network design for biodiversity and ecosystem services. *Ecological Complexity*, 26, pp.45-56.
- Ribaudo, M.O., Zilberman, D., Taylor, R. and Gordon, B., 2017. Conservation programs can accomplish more with less by improving cost-effectiveness. *Choices*, 32(4), pp.1-6.
- Rice, J., Seixas, C.S., Zaccagnini, M.E., Bedoya-Gaitán, M., Valderrama, N., Anderson, C.B., Arroyo, M.T.K., Bustamante, M., Cavender-Bares, J., Diaz-de-Leon, A. and Fennessy, S., 2018. The IPBES regional assessment report on biodiversity and ecosystem services for the Americas. Bonn, Germany.
- Ricketts, T. and Imhoff, M., 2003. Biodiversity, urban areas, and agriculture: locating priority ecoregions for conservation. *Conservation ecology*, 8(2).
- Robertson, K.R., Anderson, R.C. and Schwartz, M.W., 1997. The tallgrass prairie mosaic. In *Conservation in highly fragmented landscapes* (pp. 55-87). Boston, MA: Springer New York.
- Rocchio, J., 2007. Floristic quality assessment indices for Colorado plant communities.
- Römermann, C., Bernhardt-Römermann, M., Kleyer, M. and Poschlod, P., 2009. Substitutes for grazing in semi-natural grasslands—do mowing or mulching represent valuable alternatives to maintain vegetation structure?. *Journal of Vegetation Science*, 20(6), pp.1086-1098.
- Rowe, H.I., 2010. Tricks of the trade: techniques and opinions from 38 experts in tallgrass prairie restoration. *Restoration Ecology*, 18, pp.253-262.
- Rowe, H.I., 2013. Prairie restorations can protect remnant tallgrass prairie plant communities. *The American Midland Naturalist*, 170(1), pp.26-38.
- RStudio Team, 2020. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. <http://www.rstudio.com/>.
- Samson, F. and Knopf, F., 1994. Prairie conservation in north america. *BioScience*, 44(6), pp.418-421.
- Samson, F.B., Knopf, F.L., Gibson, D.J., Seastedt, T.R. and Briggs, J.M., 1996. Management practices in tallgrass prairie: large-and small-scale experimental effects on species composition. *Ecosystem Management: Selected Readings*, pp.106-115.
- Schwoertzig, E., Poulin, N., Hardion, L. and Trémolières, M., 2016. Plant ecological traits highlight the effects of landscape on riparian plant communities along an urban–rural gradient. *Ecological indicators*, 61, pp.568-576. <https://www.sciencedirect.com/science/article/pii/S1470160X15005439?via%3Dihub>
- Seidl, N.P. and Golobič, M., 2020. Quantitative assessment of agricultural landscape heterogeneity. *Ecological Indicators*, 112, p.106115.
- Shoo, L.P., Catterall, C.P., Beyer, H.L., Cockbain, P., Duncan, M., Robson, T., Roche, D., Taylor, H., White, Z. and Wilson, K., 2021. Smart allocation of restoration funds over space and time. *Ecological Applications*, 31(8), p.e02448.
- Sluis, W.J., Bowles, M. and Jones, M., 2018. Multiscale metrics differentiate among tallgrass prairie restorations and remnant ecosystems along a restorative continuum. *Restoration Ecology*, 26(3), pp.466-475.
- Smith, D., Williams, D., Houseal, G., Henderson, K., 2010. *The Tallgrass Prairie Center guide to prairie restoration in the Upper Midwest*. University of Iowa Press.
- Smith, D.D., 1992. Tallgrass prairie settlement: prelude to the demise of the tallgrass ecosystem. In *Proceedings of the Twelfth North American Prairie Conference* (pp. 195-199). Cedar Falls: University of Northern Iowa.
- Spyreas, G., 2019. Floristic Quality Assessment: a critique, a defense, and a primer. *Ecosphere*, 10(8).
- Steinauer, E.M. and Collins, S.L., 1996. "Prairie Ecology – The Tallgrass Prairie", in Samson, F.B. and Knopf, F.L. (ed.) *Prairie Conservation: Preserving North American's Most Endangered Ecosystem*. Washington, D.C.; Island Press, pp.39-52.

- Suding, K.N., 2011. Toward an era of restoration in ecology: successes, failures, and opportunities ahead. *Annual review of ecology, evolution, and systematics*, 42(1), pp.465-487.
- Swink, F. and G. Wilhelm, 1979. *Plants of the Chicago region*. Revised and expanded edition with keys. Lisle, Illinois: The Morton Arboretum.
- Swink, F. and G. Wilhelm, 1994. *Plants of the Chicago region*, 4th ed. Indianapolis, Indiana: Indiana Academy of Science.
- Taft, J.B., Hauser, C. and Robertson, K.R., 2006. Estimating floristic integrity in tallgrass prairie. *Biological Conservation*, 131(1), pp.42-51.
- Taft, J.B., Wilhelm, G.S., Ladd, D.M. and Masters, L.A., 1997. Floristic quality assessment for vegetation in Illinois, a method for assessing vegetation integrity (p. 29). Westville, Illinois: Illinois Native Plant Society.
- Theodorou, P., 2022. The effects of urbanisation on ecological interactions. *Current Opinion in Insect Science*, 52, p.100922.
- Theshibboleth (Creator Username), 2017. *United States Prairies.svg*. [Digital Image] Available at: https://commons.wikimedia.org/wiki/File:United_States_Prairies.svg (Accessed Apr 2025) Licensing: CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>. No changes made.
- Thomander, S.D. and Krosnick, J.A., 2024. ‘Question and Questionnaire Design’, in J.E. Edlund and A.L. Nichols (eds.) *The Cambridge Handbook of Research Methods and Statistics for the Social and Behavioral Sciences: Volume 2: Performing Research*. Cambridge: Cambridge University Press (Cambridge Handbooks in Psychology), pp. 352–370.
- Török, P., Brudvig, L.A., Kollmann, J., N Price, J. and Tóthmérész, B., 2021. The present and future of grassland restoration. *Restoration Ecology*, 29, p.e13378.
- Towey, B. J., 2025. Email to Brian Towey, 29 January.
- United States Department of Agriculture, 2024. “United States – Crop Production Maps”, USDA Foreign Agricultural Service. Available at: https://ipad.fas.usda.gov/rssiws/al/us_cropprod.aspx (Accessed May 2025)
- United States Geological Survey, 2024. Annual National Land Cover Database (NLCD) Collection 1 Products. [Dataset] DOI: 10.5066/P94UXNTS Available at: <https://www.sciencebase.gov/catalog/item/655ceb8ad34ee4b6e05cc51a> (Accessed: Nov 2024)
- Van Horn, G., Mac Aodha, O., Song, Y., Cui, Y., Sun, C., Shepard, A., Adam, H., Perona, P. and Belongie, S., 2018. The inaturalist species classification and detection dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8769-8778).
- Wagle, P. and Gowda, P.H., 2018. Tallgrass prairie responses to management practices and disturbances: A review. *Agronomy*, 8(12), p.300.
- Wagle, P., Gowda, P.H., Northup, B.K., Starks, P.J. and Neel, J.P., 2019. Response of tallgrass prairie to management in the US Southern great plains: Site descriptions, management practices, and eddy covariance instrumentation for a Long-Term Experiment. *Remote Sensing*, 11(17), p.1988.
- Weaver, J.E. and Fitzpatrick, T.J., 1932. Ecology and relative importance of the dominants of tall-grass prairie. *Botanical Gazette*, 93(2), pp.113-150.
- Weaver, J.E., 1927. Some ecological aspects of agriculture in the prairie. *Ecology*, 8(1), pp.1-17.
- Wetli, P., 2023. “Death of a Prairie: Bulldozers Arrived at Bell Bowl at Dawn”, WTTW News, Mar. 9, Available at: <https://news.wttw.com/2023/03/09/death-prairie-bulldozers-arrived-bell-bowl-dawn>
- Wickham H, 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.
- Wilby, R.L. and Perry, G.L., 2006. Climate change, biodiversity and the urban environment: a critical review based on London, UK. *Progress in physical geography*, 30(1), pp.73-98.
- Wilhelm, G. and Rericha, L., 2017. *Flora of the Chicago Region: A Floristic and Ecological Synthesis*, Indianapolis, Indiana: Indiana Academy of Science.
- Williams, J.R. and Diebel, P.L., 1996. “The Economic Value of the Prairie”, in in Samson, F.B. and Knopf, F.L (ed.) *Prairie Conservation: Preserving North American’s Most Endangered Ecosystem*. Washington, D.C.; Island Press, pp.19-35.
- Williams, N.S., Morgan, J.W., McDonnell, M.J. and McCarthy, M.A., 2005. Plant traits and local extinctions in natural grasslands along an urban–rural gradient. *Journal of Ecology*, 93(6), pp.1203-1213.

- Williams, S.L., 2005. Growth and development of the conservation profession: a matter of degrees. *Collections*, 2(2), pp.83-93.
- Yu, Z. and Lu, C., 2018. Historical cropland expansion and abandonment in the continental US during 1850 to 2016. *Global Ecology and Biogeography*, 27(3), pp.322-333.
- Zirbel, C.R., Grman, E., Bassett, T. and Brudvig, L.A., 2019. Landscape context explains ecosystem multifunctionality in restored grasslands better than plant diversity. *Ecology*, 100(4), p.e02634.
- Zuur, A., Ieno, E., Walker, N., Saveliev, A., Smith, G., 2009. *Mixed Effects Models and Extensions in Ecology with R*. New York City: Springer Science+Business Media

Appendix Information

Appendix – Enter the Landscape Matrix: Effects of Anthropogenic Landscapes and Land Management on the Floral Quality of Tallgrass Prairies in Illinois, USA

Andy Sima



Sunset over Glacial Park Conservation Area south prairie.

Appendix Table of Contents

Appendix A: Full Formulae and All Calculated Variables	45
Table A1. Full Formulae of Response Variables	45
Table A2. GLM Summary	46
Table A3. Description of Explored Variables	48
In Defense of FQI	52
Appendix B: Extra Figures	53
Appendix C: Sample Management Survey	65
Appendix D: Full Species List	75
Appendix E: Information Regarding All Zones and Sites	82
Index Table of Zones, Sites, and Ownership	83
Index Table of Zone Scores, Ranked by Native FQI	84
Zone 1 (Site 1): Hinsdale INAI Prairie	85
.....	
Zone 30 (Site 30): Bridgeview INAI Prairie	143
Appendix F: Full Coding Scripts	145
Python-Based Land Cover Analysis	145
Python-Based Data Cleaning and Measuring	161
R-Based Statistical Analysis and Data Exploration	193

Appendix A: Full Formulae and All Calculated Variables

Table A1. Full Formulae of Response Variables

Table A1. Full formulae of all calculated response variables, including unreported component variables. Used as per-plot metrics used for statistical analysis. Once a value was calculated for all ten plots in a zone, they were averaged to get a value representative of the zone itself, excepting Shannon, Simpson, and Gamma Diversity (and their component metrics), which could only be calculated on a zone-wide basis. Pest encroachment metrics were omitted due to a lack of proper citations.

I used a modified version of Shannon's diversity utilizing Relative Importance Values (IV 200) instead of pure species counts, allowing me to calculate Shannon's diversity from mean cover and frequency. Relative Importance Values represent a single measurement of a species' presence within a community (Taft, Hauser, and Robertson 2006). Because it required both mean frequency and coverage, Shannon's diversity was technically only calculated on a zone-wide basis, automatically considering all ten plots per zone but not averaging them.

Where i represents an individual from a given inventory (*nat* for native, *all* for all, and *int* for introduced species), Ci represents the C-value of an individual, $Covi.n$ is the number of instances of cover class of each individual species while n represents the cover class itself (classes range from 1 through 7, see Table 1; the sum of instances of cover class times their midpoint), $mp.n$ is the midpoint of each cover class, *total* represents the total number of samples that a species could possibly appear in (10 per zone), *frequi* is the frequency of an individual specie's appearance in replicate plots (i.e. how many times was this species sampled out of ten replicate plots), and *zone* represents the summation across the entire zone, i.e. all ten plots. A sum of an i individual term (e.g. $\sum Mcovi$) represents the summation of that term for all individual species.

** Denotes a metric that was reported and represents one of the primary response variables measured during statistical analysis. All other formulae were required but unreported or ultimately unused in analysis, but previously calculated.

(*Native/All*) Denotes a metric that can be calculated for both native species only or all species depending on the subset of species utilized in calculation. However, some formulae (e.g. Mean-C) differ between native and all calculations, and thus retain their own formula.

Parameter	Formula	Source
Species Richness (Native)**	$N = \sum i(nat)$	Rocchio 2007
Species Richness (All)	$S = \sum i(all)$	Rocchio 2007
Species Richness (Introduced)	$Si = \sum i(int)$	Rocchio 2007
Percentage Introduced Species**	$Perc(int) = \left(\frac{\sum i(int)}{S} \right) * 100$	Rocchio 2007

Mean-C (Native)	$MeanCn = \frac{\sum Ci(nat)}{N}$	Rocchio 2007
Mean-C (All)	$MeanC(all) = \frac{\sum Ci(all)}{S}$	Rocchio 2007
Mean-C Difference	$MeanCDif = MeanC(all) - MeanC(nat)$	Rocchio 2007
Median-C (Native)	$MedC(nat) = Median(Ci(nat))$	Fiaschi et al. 2023
Median-C (All)	$MedC(all) = Median(Ci(all))$	Fiaschi et al. 2023
Median-C Difference	$MedCDif = MedC(all) - MedC(nat)$	Fiaschi et al. 2023; Rocchio 2007

Native Floristic Quality Index (FQI)**	$FQI = MeanCn * \sqrt{N}$	Taft, Hauser, and Robertson 2006
--	---------------------------	----------------------------------

Parameter	Formula	Source
Floristic Quality Index (All)	$FQI(all) = MeanC(all) * \sqrt{N}$	Taft, Hauser, and Robertson 2006
FQI Difference	$FQIdif = FQI(all) - FQI$	Rocchio 2007; Taft, Hauser, and Robertson 2006
Adjusted Floristic Quality Index (AFQI)**	$AFQI = \left(\frac{MeanCn}{10} * \frac{\sqrt{N}}{\sqrt{S}} \right) * 100$	Rocchio 2007

Mean Cover per Species	$MCovi = \frac{\sum_{n=1}^7 ((\sum Covi.n) * mp.n)}{total}$	Rocchio 2007; Taft, Hauser, and Robertson 2006
Relative Coverage (Native/All)	$RCovi = \frac{MCovi}{\sum MCovi}$	Taft, Hauser, and Robertson 2006
Relative Frequency (Native/All)	$RFreqi = \frac{freqi}{\sum freqi}$	Taft, Hauser, and Robertson 2006
Relative Importance Value (IV 200) (Native/All)	$p = \frac{RCovi + RFreqi}{\sum (RCovi + RFreqi)}$	Taft, Hauser, and Robertson 2006
Shannon-Wiener Diversity Index (Native/All**)	$H' = - \sum p * \ln(p)$	Taft, Hauser, and Robertson 2006
Simpson's Diversity Index (Native/All)	$D = 1 - \sum p^2$	Taft, Hauser, and Robertson 2006

Native Gamma Diversity**	$\gamma = \sum i(nat).zone$	Sluis, Bowles, and Jones 2018
All Gamma Diversity	$\gamma = \sum i(all).zone$	Sluis, Bowles, and Jones 2018

Table A2. GLM Summary

Table A2. Full statistical values of the Generalized Linear Model discussed in the results and conclusion of this report. Contains summary information for all five primary response variables plus native gamma diversity. N=23 for all. All are based on GLM models combining the ratio of rural to suburban (log-transformed), the site budget per hectare of zone, and the sum of handweeding intensity as drawn from the management survey.

Native Species Richness (family = Gaussian, link = log)

	Estimate	Std. Error	T-Value (Wald)	P-Value
(Intercept)	2.14	0.12	17.79	2.64e-13
Rural-Suburban Ratio (Log)	0.048	0.022	2.15	0.045
Site Budget per Hect. Zone	8.80e-06	3.932e-06	2.24	0.037
Hand weeding Intensity Sum	0.016	0.014	1.15	0.26

Percent Introduced Species (family = Gamma, link = log)

	Estimate	Std. Error	T-Value (Wald)	P-Value
(Intercept)	3.03	0.24	12.87	7.93e-11
Rural-Suburban Ratio (Log)	0.052	0.043	1.21	0.24
Site Budget per Hect. Zone	-2.74e-05	1.08e-05	-2.52	0.021
Hand weeding Intensity Sum	-0.028	0.029	-0.96	0.35

Native FQI (family = Gaussian, link = log)

	Estimate	Std. Error	T-Value (Wald)	P-Value
(Intercept)	2.44	0.11	22.92	2.65e-15
Rural-Suburban Ratio (Log)	0.049	0.019	2.57	0.019
Site Budget per Hect. Zone	8.89e-06	3.30e-06	2.70	0.014
Hand weeding Intensity Sum	0.023	0.012	1.87	0.077

AFQI (family = Gaussian, link = log)

	Estimate	Std. Error	T-Value (Wald)	P-Value
(Intercept)	3.57	0.064	55.46	<2e-16
Rural-Suburban Ratio (Log)	0.020	0.011	1.83	0.082
Site Budget per Hect. Zone	5.46e-06	2.25e-06	2.42	0.026
Hand weeding Intensity Sum	0.018	0.0074	2.40	0.027

Shannon Diversity (family = Gamma, link = log)

	Estimate	Std. Error	T-Value (Wald)	P-Value
(Intercept)	1.13	0.026	42.77	<2e-16
Rural-Suburban Ratio (Log)	0.012	0.0048	2.48	0.023
Site Budget per Hect. Zone	8.24e-07	1.21e-06	0.68	0.51
Hand weeding Intensity Sum	0.0049	0.0032	1.53	0.14

Native Gamma Diversity (family = Poisson, link = log)

	Estimate	Std. Error	Z-Value (Wald)	P-Value
(Intercept)	3.34	0.067	50.07	<2e-16
Rural-Suburban Ratio (Log)	0.0065	0.012	0.56	0.57

Site Budget per Hect. Zone	4.90e-06	2.70e-06	1.82	0.069
Hand weeding Intensity Sum	0.018	0.0078	2.35	0.019

Table A3. Description of Explored Variables

Table A3. A brief summary and description of the majority of the response and predictor variables that were calculated and explored for this study. This includes variables that were statistically analysed for trends but were not, ultimately, reported on in any way. It is largely safe to assume that if a predictor variable is omitted from the report, it means that there was no significant correlation between it and any of the reported response variables. Omitted response variables were unreported largely because of redundancy to the reported response variables. Further variables omitted here, such as woody encroachment metrics, were calculated but were not analysed to a meaningful degree.

Where (Native/All) represents that a variable can be calculated for both native only or all species, depending on the subset of species analysed.

* Denotes variables that were reported on in the paper.

VARIABLE NAME	DESCRIPTION
---------------	-------------

RESPONSE VARIABLES	
Species Richness (Native*/All)	Number of unique species, averaged across 10 plots.
Percentage of Introduced Species*	Percentage of species that are introduced/nonnative out of the total pool of species,
FQI (Native*/All)	Floristic Quality Index, roughly a mean of all coefficients of conservatism, represents a measure of a plant community's quality, averaged across 10 plots.
AFQI*	FQI adjusted to include introduced species as a natural part of the metric, averaged across 10 plots.
Shannon Diversity Index (Native/All*)	A measure of diversity that considers both the number of unique species and their relative abundance, measured at the zone scale.
Mean-C (Native/All)	The average of all coefficients of conservatism, roughly measures quality, averaged across 10 plots.
Simpson Diversity Index (Native/All)	Similar to Shannon diversity, but more accurately described as a measure of species evenness, measures at the zone scale.
Gamma Diversity (Native*/All)	Number of unique species inventoried across all 10 plots, not averaged.

PREDICTOR VARIABLES	
Zone Area (Hectares)*	The total area of just the zone.
Site Area (Hectares)	The total area of the entire site, zone included.
Binary Land Classification*	Whether a zone/site is rural or suburban, based on which land class is greater at a 2000 m buffer.

Distance to Chicago (km)	How far the geometric center of the zone is from the geometric center of Chicago (roughly the Heart of Chicago neighborhood).
Average Percentage of Total Wilderness*	The average percentage of all NLCD wilderness classes across eight buffer distances.
Rate of Change of Percentage of Total Wilderness	The slope of the percentage of NLCD wilderness classes as plotted against eight buffer distances.
Relative Rate of Change of Percentage of Total Wilderness	Rate of change as above, but divided by the average of the class.
Average Percentage of Total Open Wilderness (Grassland)*	The average percentage of NLCD open wilderness classes (grassland, non-forest) across eight buffer distances.
Rate of Change of Percentage of Open Wilderness (Grassland)	The slope of the percentage of NLCD open wilderness classes as plotted against eight buffer distances.
Relative Rate of Change of Percentage of Open Wilderness (Grassland)	Rate of change as above, but divided by the average of the class.
Average Percentage of Total Rural Matrix*	The average percentage of all NLCD agricultural classes across eight buffer distances.
Rate of Change of Percentage of Rural Matrix	The slope of the percentage of NLCD agricultural classes as plotted against eight buffer distances.
Relative Rate of Change of Percentage of Rural Matrix	Rate of change as above, but divided by the average of the class.
Average Percentage of Suburban Matrix*	The average percentage of all NLCD development classes across eight buffer distances.
Rate of Change of Percentage of Suburban Matrix	The slope of the percentage of NLCD development classes as plotted against eight buffer distances.
Relative Rate of Change of Percentage of Suburban Matrix	Rate of change as above, but divided by the average of the class.
Percentage of Total Wilderness at 50 m Buffer	All NLCD wilderness classes as a percent of total buffer area at a 50 m buffer.
Percentage of Open Wilderness (Grassland) at 50 m Buffer	All NLCD open wilderness classes as a percent of total buffer area at a 50 m buffer.
Percentage of Rural Matrix at 50 m Buffer	All NLCD agricultural classes as a percent of total buffer area at a 50 m buffer.
Percentage of Suburban Matrix at 50 m Buffer	All NLCD developed classes as a percent of total buffer area at a 50 m buffer.
Ratio of Rural Matrix-Suburban Matrix at 50 m Buffer*	The ratio of rural percentage over suburban percentage at a 50 m buffer.
Percentage of Total Wilderness at 1500 m Buffer	All NLCD wilderness classes as a percent of total buffer area at a 1500 m buffer.
Percentage of Open Wilderness (Grassland) at 1500 m Buffer	All NLCD open wilderness classes as a percent of total buffer area at a 1500 m buffer.
Percentage of Rural Matrix at 1500 m Buffer	All NLCD agricultural classes as a percent of total buffer area at a 1500 m buffer.

Percentage of Suburban Matrix at 1500 m Buffer	All NLCD developed classes as a percent of total buffer area at a 1500 m buffer.
Ratio of Rural Matrix-Suburban Matrix at 1500 m Buffer*	The ratio of rural percentage over suburban percentage at a 1500 m buffer.
Percentage of Total Wilderness at 3000 m Buffer	All NLCD wilderness classes as a percent of total buffer area at a 3000 m buffer.
Percentage of Open Wilderness (Grassland) at 3000 m Buffer	All NLCD open wilderness classes as a percent of total buffer area at a 3000 m buffer.
Percentage of Rural Matrix at 3000 m Buffer*	All NLCD agricultural classes as a percent of total buffer area at a 3000 m buffer.
Percentage of Suburban Matrix at 3000 m Buffer	All NLCD developed classes as a percent of total buffer area at a 3000 m buffer.
Ratio of Rural Matrix-Suburban Matrix at 3000 m Buffer	The ratio of rural percentage over suburban percentage at a 3000 m buffer.
Log-Ratio of Rural Matrix-Suburban Matrix at 2000 m SITE Buffer	The ratio of rural percentage over suburban percentage at a 2000 m buffer from the edge of the site. All other buffers use the edge of the zone as their bases.
Average Log-Ratio of Total Wilderness to Dominant Class	The log-transformed ratio of the average percentage of total wilderness to the dominant (rural or suburban) class.
Average Log-Ratio of Open Wilderness to Closed Wilderness	The log-transformed ratio of the average percentage of open wilderness (grassland) to closed wilderness (forest)
Average Log-Ratio of Rural Matrix-Suburban Matrix*	The log-transformed ratio of the average percentage of rural classes to suburban classes.
Log-Ratio of Rural Matrix-Suburban Matrix at 50 m Buffer log	The log-transformed ratio of the percentage of rural classes to suburban classes at 50 m buffer.
Log-Ratio of Rural Matrix-Suburban Matrix at 1500 m Buffer log	The log-transformed ratio of the percentage of rural classes to suburban classes at 1500 m buffer.
Log-Ratio of Rural Matrix-Suburban Matrix at 3000 m Buffer log	The log-transformed ratio of the percentage of rural classes to suburban classes at 3000 m buffer.
Age of Restoration – Grouped*	The age of a zone, grouped into three categories to account for unknown remnant ages. Collected from management survey.
Restoration to Remnant Scale	On a scale of 0 (totally restoration) to 10 (totally remnant), how much remnant versus restoration is part of a zone, as reported by managers. Collected from management survey.
Age of Restoration	The number of years since a restoration began, not including remnants. Collected from management survey.
Age of Restoration with Remnants	The number of years since a restoration began, with 1837 (187 years; development of the steel plow) as the age set for remnants. Collected from management survey.

Number of Methods Used	Manager-reported number of management techniques utilized in a zone. Options include burning, mowing, hand weeding, herbicide, seeding, and other. Collected from management survey.
Number of Methods Used – Grouped*	Same as above, but grouped based on number of methods used. Collected from management survey.
Mowing – Yes/No	Presence or absence of mowing as a management tool. Collected from management survey.
Seeding – Yes/No	Presence or absence of seeding as a management tool. Collected from management survey.
Hand Weeding – Yes/No*	Presence or absence of hand weeding as a management tool. Collected from management survey.
Permanent Staff Number	Number of permanently-employed, full-time staff working on a zone/site. Collected from management survey.
Management Intensity*	Manager-reported estimated intensity of management applied to a zone, from a scale of 0 (no management) to 10 (very high management). Collected from management survey.
Burning Sum	Rough summation of responses for questions related to the intensity of burning as a management tool. Collected from management survey.
Mowing Sum	Rough summation of responses for questions related to the intensity of mowing as a management tool. Collected from management survey.
Seeding Sum	Rough summation of responses for questions related to the intensity of seeding as a management tool. Collected from management survey.
Herbicide Sum	Rough summation of responses for questions related to the intensity of herbicide application as a management tool. Collected from management survey.
Handweed Sum	Rough summation of responses for questions related to the intensity of hand weeding as a management tool. Collected from management survey.
Staff Sum*	Summation of responses for questions related to permanent staff, part-time staff, volunteers, and contract labor. Collected from management survey.
Management Sum*	Summation of responses for all questions related to the application and intensity of management, without consideration of number of staff. Collected from management survey.

Management Sum with Staff Sum	Same as above, but with sum of staff as a further modifier. Collected from management survey.
Management Sum with Staff Sum Per Hectare of Zone	Same as above, but normalized against the size of the zone. Collected from management survey.
Management Sum with Staff Sum per Hectare of Site	Same as above, but normalized against the size of the site. Collected from management survey.
Funds per Zone	Manager-reported estimation of the money spent on the management of a zone. Collected from management survey.
Funds per Zone Grouped	Manager-reported estimation of the money spent on the management of a zone, grouped into distinct categories. Collected from management survey.
Operating Organization Estimated Budget	Manager-reported estimation of the annual operating budget of the organization responsible for the ownership of a zone/site. Collected from management survey.
Operating Organization Estimated Budget Grouped*	Manager-reported estimation of the annual operating budget of the organization responsible for the ownership of a zone/site, grouped into distinct categories. Collected from management survey.
Operating Organization Estimated Budget Per Managed Site	Estimated organizational budget, but divided by the estimated number of other, non-sampled sites also within the organization's ownership. Collected from management survey.
Operating Organization Estimated Budget per Hectare of Zone*	Estimated organizational budget, but divided by the size of the zone. Collected from management survey.
Operating Organization Estimated Budget Per Hectare of Site	Estimated organizational budget, but divided by the size of the site. Collected from management survey.
Public-Private Sector Type*	Whether a site is publicly-owned or privately-owned. Drawn from the Prairie State Conservation Coalition.
Special Legal Designations – Yes/No*	Whether a site has any special legal designations, such as National Wildlife Refuge or State Nature Preserve. Collected from management survey.

In Defense of FQI

While there are some recent studies that cast doubt upon the reliability of FQI and Mean C as an accurate measurement of quality (Landi and Chiarucci 2010, Fiaschi et al 2023), Mean C and FQI are still, even within studies that criticize it, cited as some of the easiest means to assess floral quality. FQI especially is a fairly strong measurement as well, considered by many American ecologists to be a reliable measure of floristic quality and robust to certain errors, such as occasional species misidentification (Spyreas 2019). They are still some of the most commonly used metrics within floral studies in the United States (Rocchio 2007; Spyreas 2019). Beyond that, I analysed several other floral quality measurements beyond Mean C and FQI, which also demonstrated consistent trends. While these limitations do have to be considered when looking at the big picture surrounding this research, I do not think that they would radically alter the results found herein, nor would they discount these conclusions entirely.

Appendix B: Extra Figures

The following figures represent maps, data, and graphs that were produced in the process of data analysis for this project, but were deemed largely unnecessary to the main body of the paper. They exist here to serve as further explanation and visualization of results and methodologies.

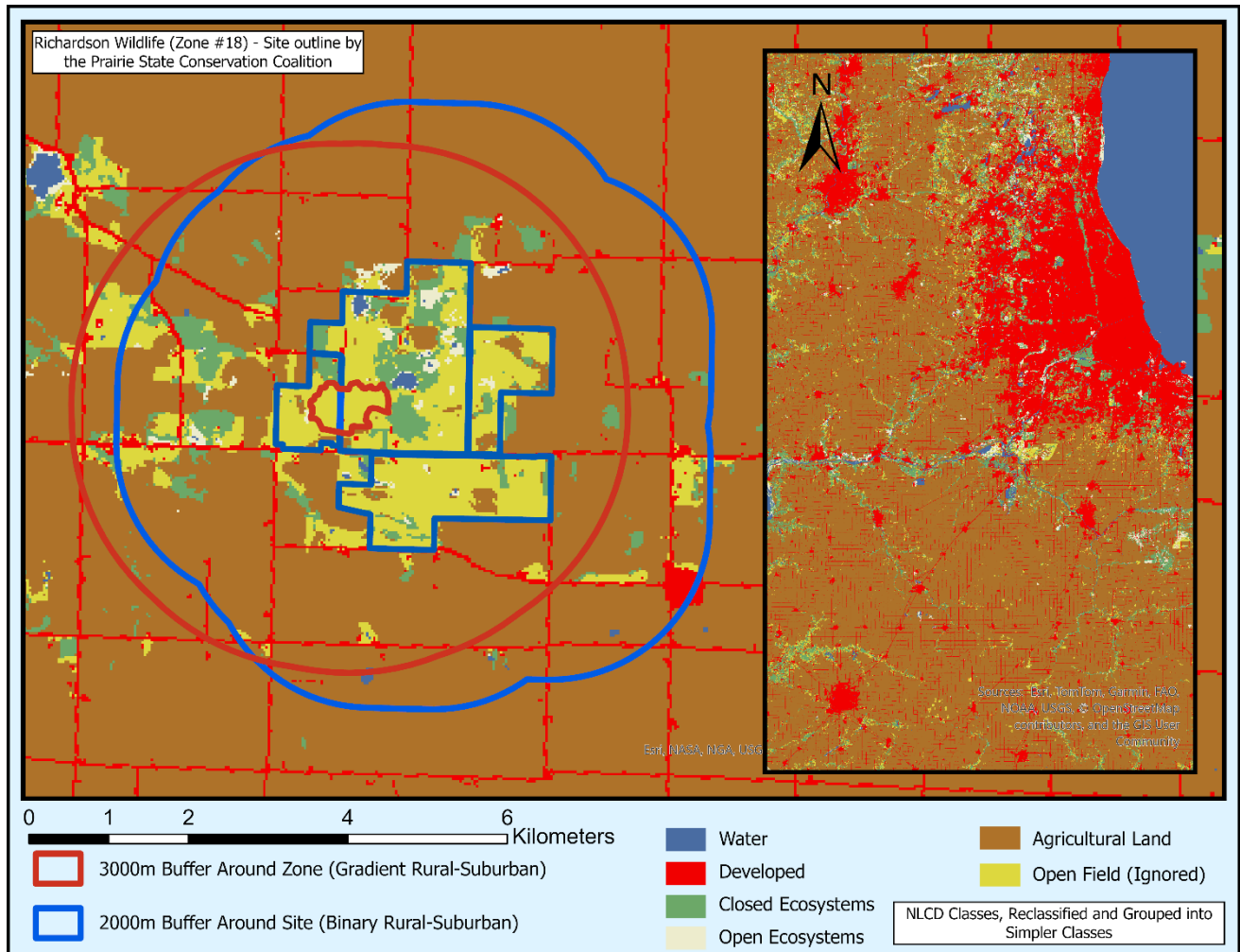


Figure B1. A map demonstrating the NLCD classes after they have been reclassified and regrouped per the land cover analysis methodology in section 2.2a. This process simplified the classes of the NLCD down to six groups. You can see that the Open Field category, which is agricultural by NLCD standards, is misidentified prairie. In reality, there is no agricultural space within the zone, but the NLCD categorizes it here as agricultural land. I ignored that category to not introduce error and misidentification.

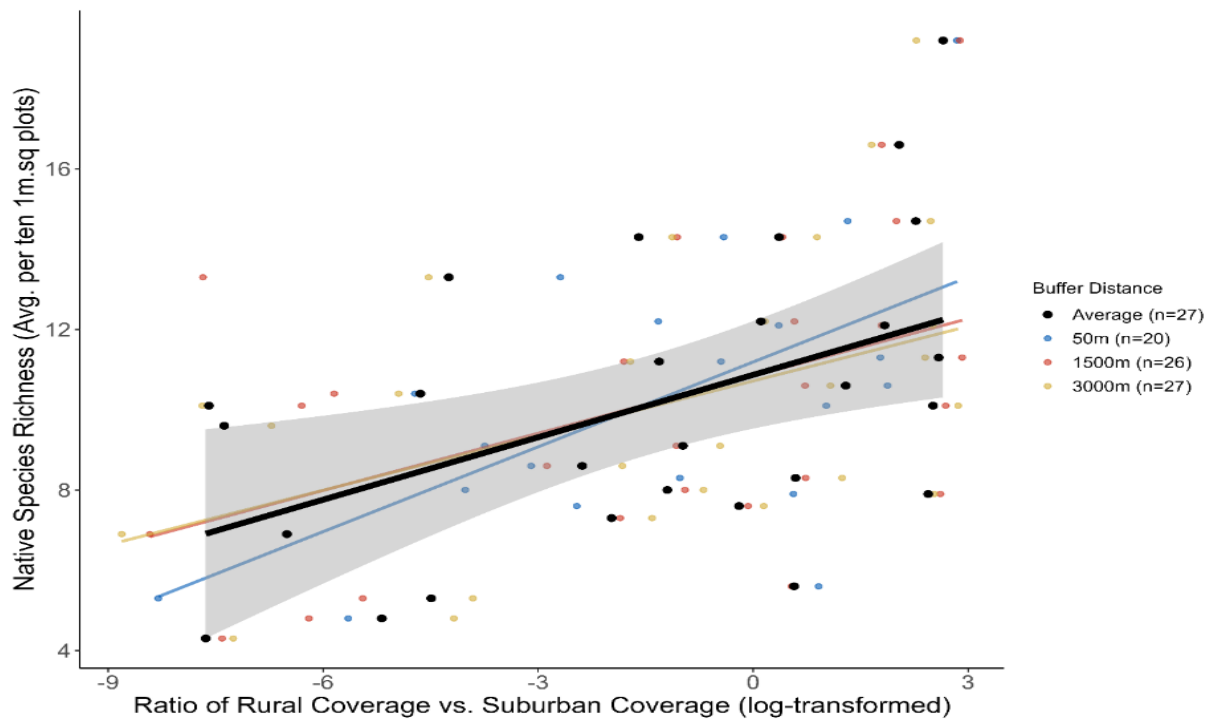


Figure B2. A composite line graph depicting native species richness of all sites against four measurements of the log-transformed average ratio of rural to suburban landscape. These measurements include the average of all eight ratios, plus the pure values of the log-transformed ratios at 50m, 1500m, and 3000m. The number (n) of points differ as some ratios were zero, which cannot be log-transformed.

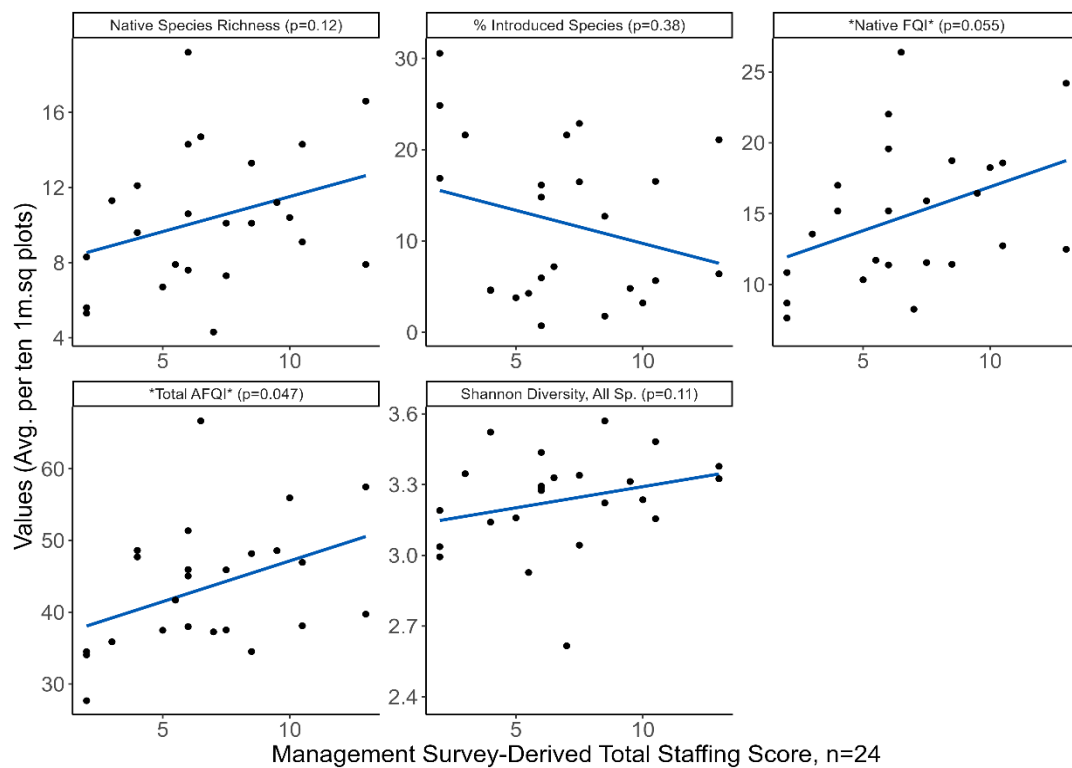


Figure B3. A series of bar graphs demonstrating the five response variables against the sum of staffing as measured from the management survey. This takes into account the number of permanent staff, seasonal staff, volunteers, and contractors as a general sense of total labor available to a given site/zone. Trends were not present in analyses of the individual staffing components on their own.

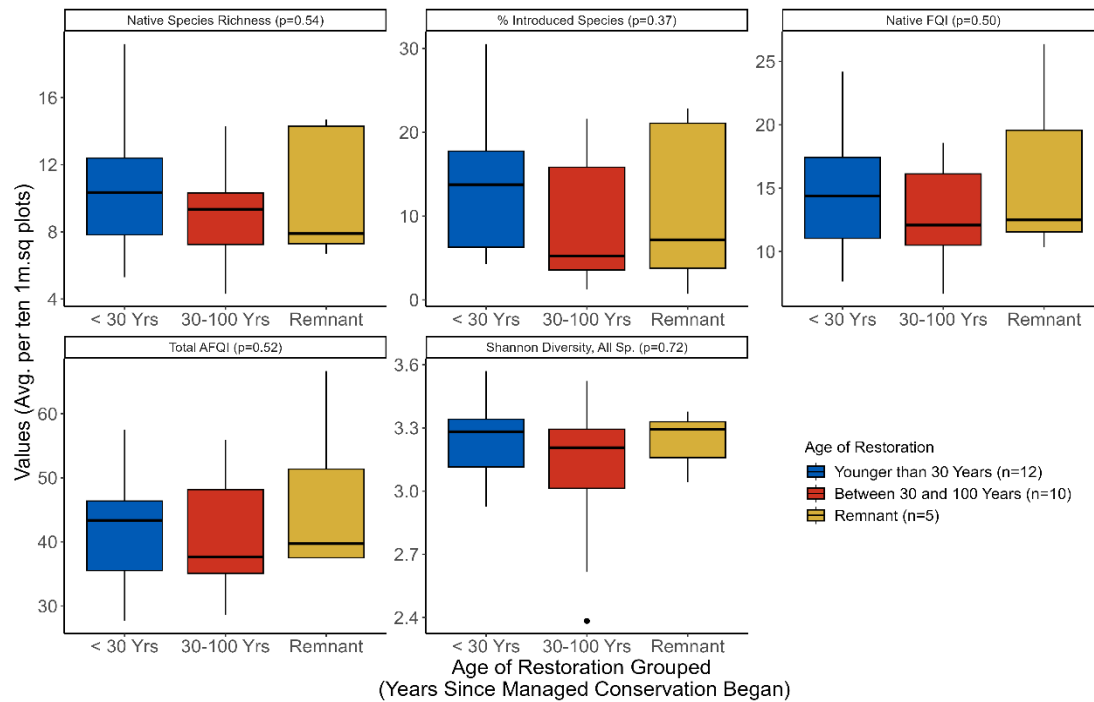


Figure B4. Five bar graphs demonstrating the response variables when compared against the age of prairies. As remnant prairies have no set age, they were grouped into a distinct category and compared against groupings of restorations, based on age. There were no significant differences between any of the classes based on age.

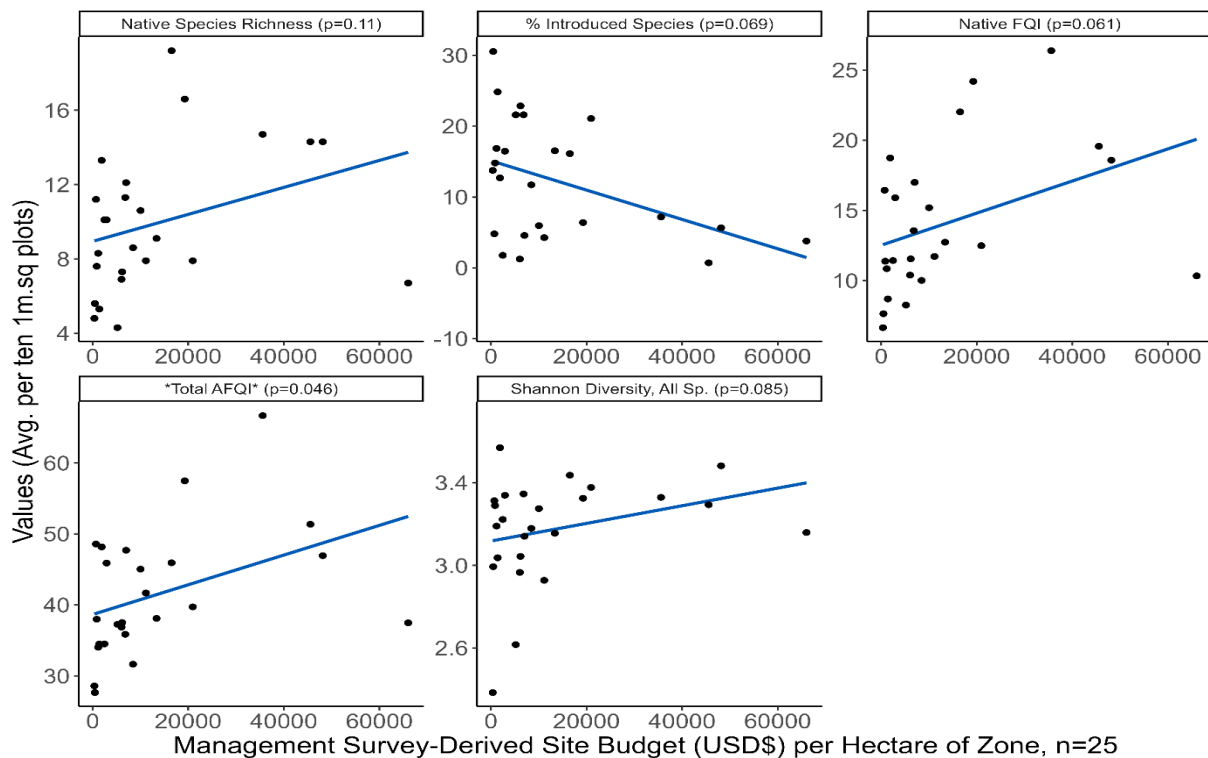


Figure B5. Line graphs demonstrating the five response variables when plotted against the survey-derived site budget when normalized against the size of just the zone, in hectares. As the budget available for each hectare of zone increases, so too does quality generally increase. However, when compared to the area of the entire site (not just the zone), this trend disappears. All trends disappeared and there was no correlation when larger operating organization budgets were normalized against the number of unsampled sites that each organization may be responsible for

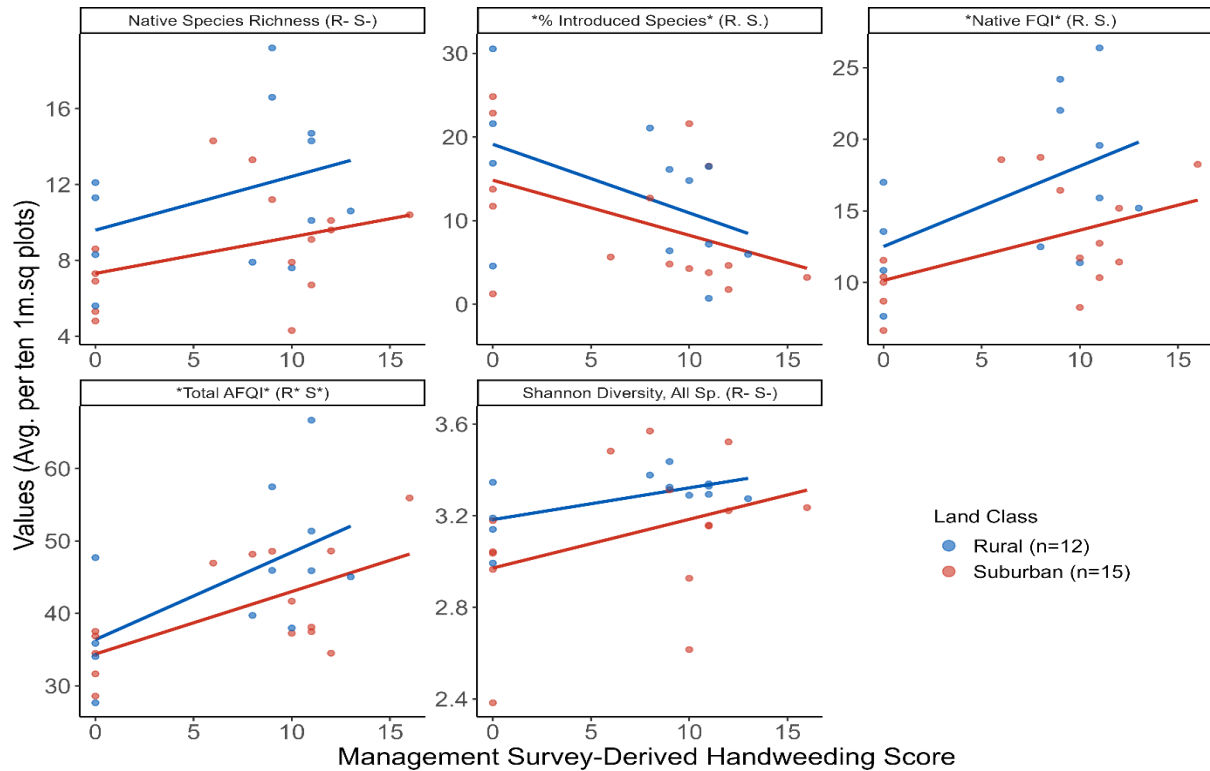


Figure B6. Line graphs depicting the five response variables when compared to the survey-derived handweeding score, but responses are grouped based upon their landscape matrix status according to rural or suburban. R stands for Rural, and S stands for Suburban. An * asterisk next to a letter represents significance, a - dash represents no significance, and a . dot represents near-significance.

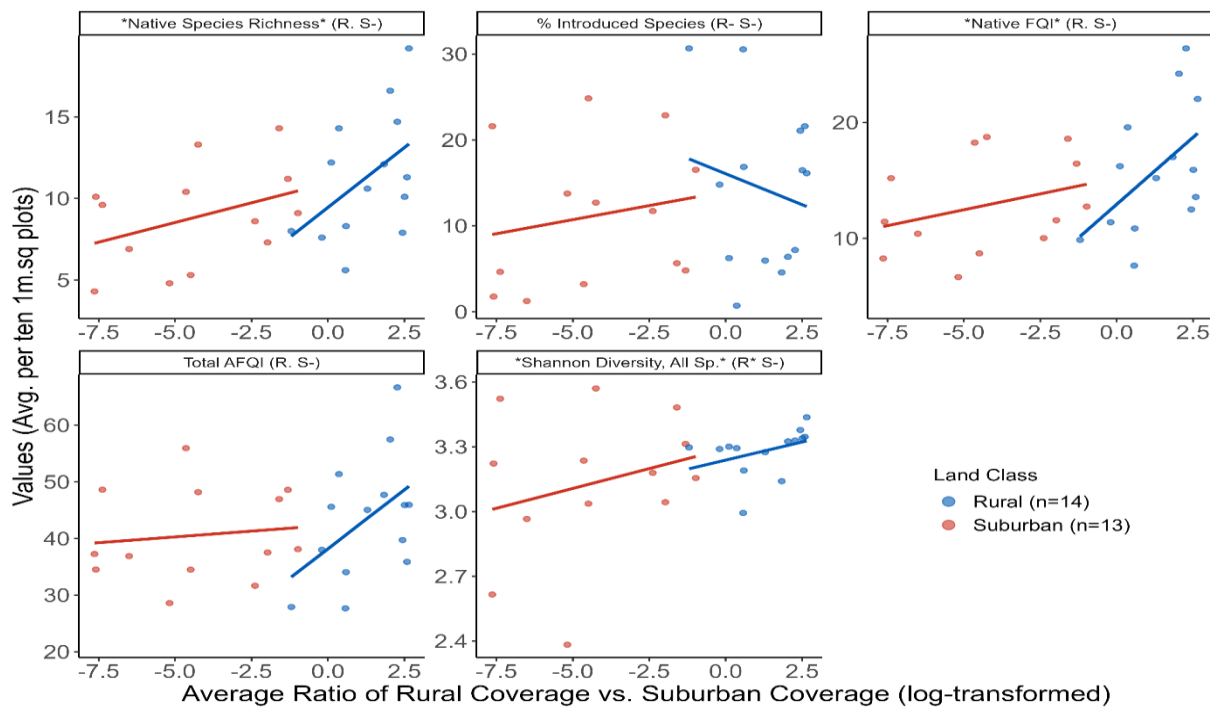


Figure B7. Line graphs depicting the five response variables when compared to the log-transformed average ratio of rural to suburban, but responses are grouped based upon their landscape matrix status according to rural or suburban. This demonstrates a more similar response among rural areas than among suburban ones.

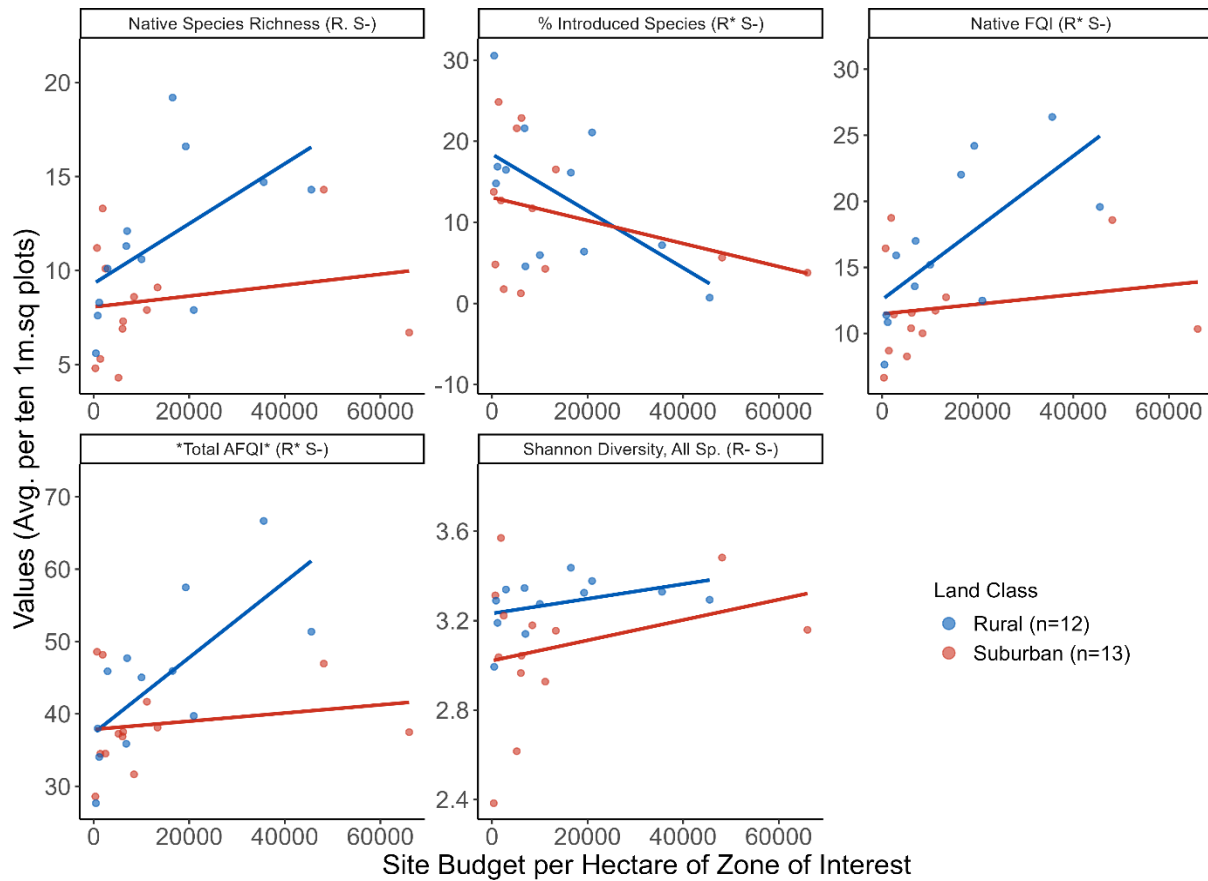


Figure B8. Line graphs depicting the five response variables when compared to the site budget per hectare of zone of interest, but responses are grouped based upon their landscape matrix status according to rural or suburban. This demonstrates a more similar response among rural areas than among suburban ones, but also demonstrates higher baseline quality in rural zones, and an increased magnitude of response among rural prairies.

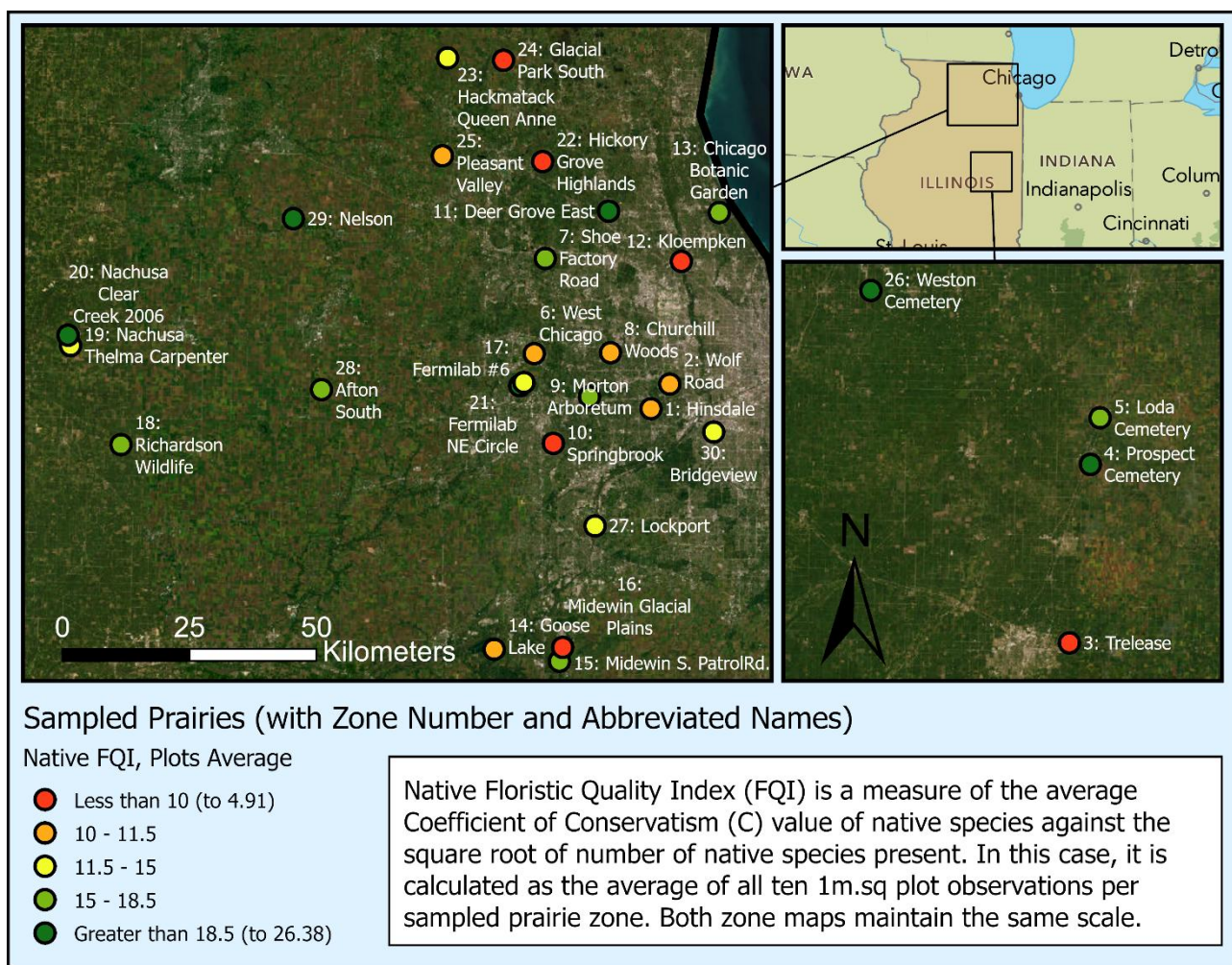


Figure B9. A results maps demonstrating the spread of sampled zones as well as their respective native species FQI scores. Although there were correlations between zone quality and anthropogenic landscape matrix, it is not necessarily clear which zones are rural or suburban based on their FQI scores alone. The spread and variability of data makes it exceedingly hard to predict whether a zone is rural or suburban. And, conversely, it is likely impossible to estimate a prairie's quality based solely on its surrounding landscape matrix.

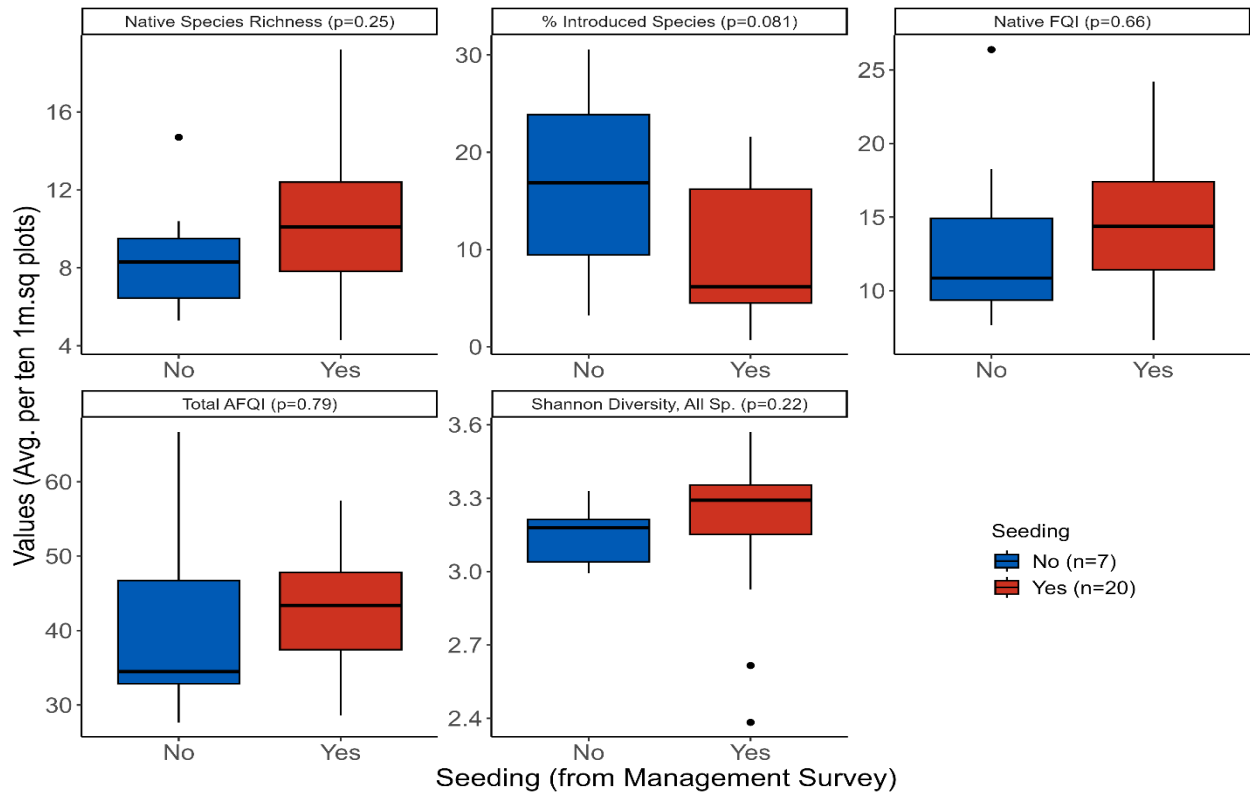


Figure B10. Box charts depicting the five response variables when compared to the presence/absence of seeding as a management tool. This demonstrates that other management techniques did not have as notable a correlation as hand weeding had, but also demonstrates a slight increase in quality with the presence/absence of other techniques.

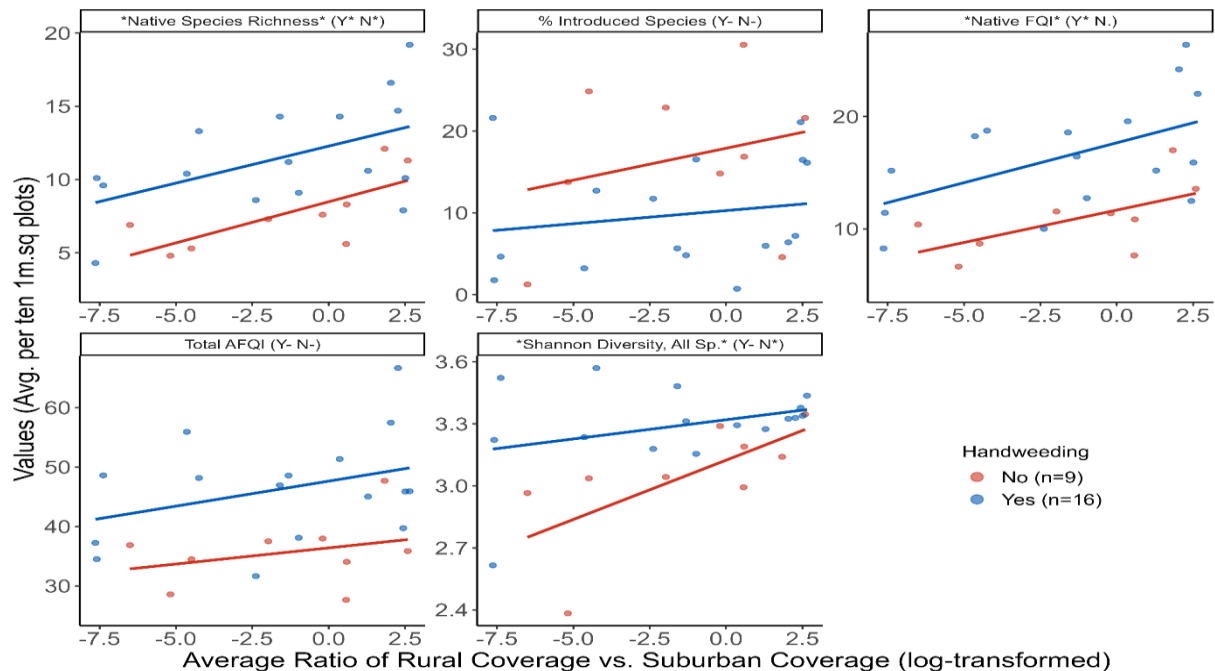


Figure B11. Line graphs depicting the five response variables when compared to the ratio of rural to suburban coverage, but responses are grouped based upon the presence or absence of hand weeding as a management tool. This demonstrates consistency in the benefits granted by the presence of handweeding. Sites with handweeding as a management tool will demonstrate a higher baseline quality than sites without, while still following trends established by other variables.

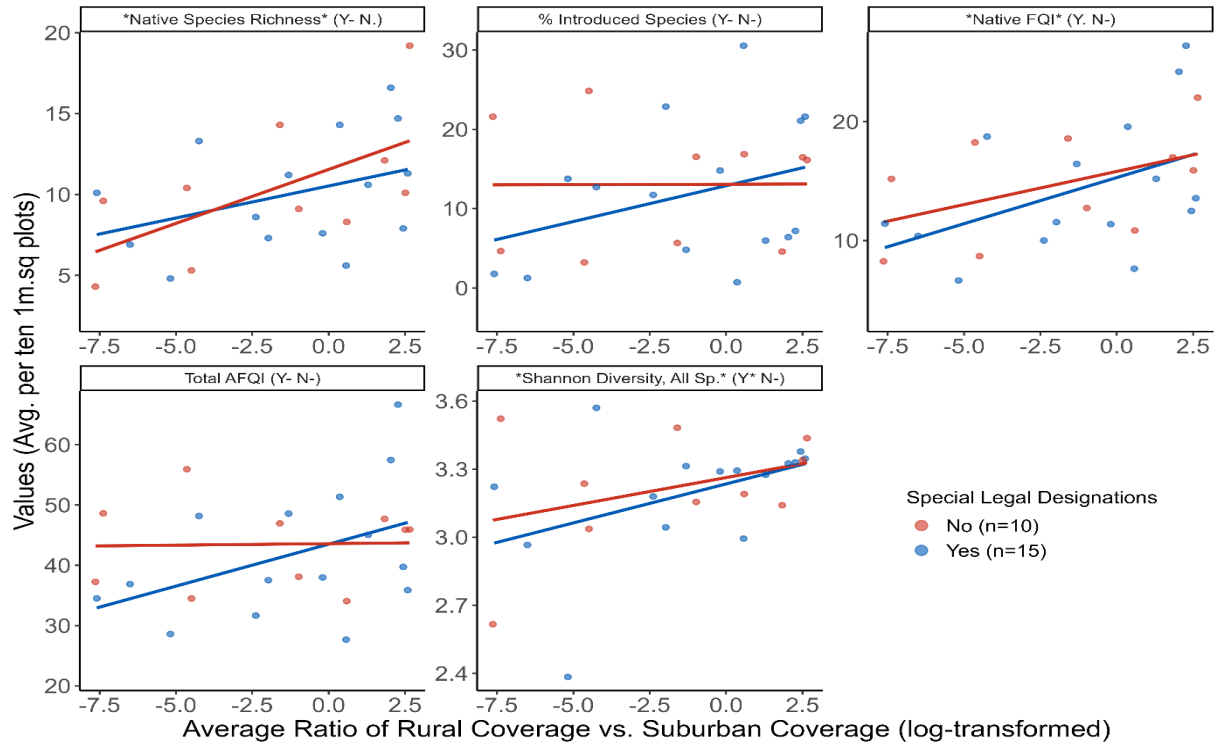


Figure B12. Line graphs depicting the five response variables when compared to the ratio of rural to suburban coverage, but responses are grouped based upon the presence or absence of special legal designations. This demonstrates consistency in the fact that legal designation has zero impact on the overall quality of a site, while still maintaining relevant trends.

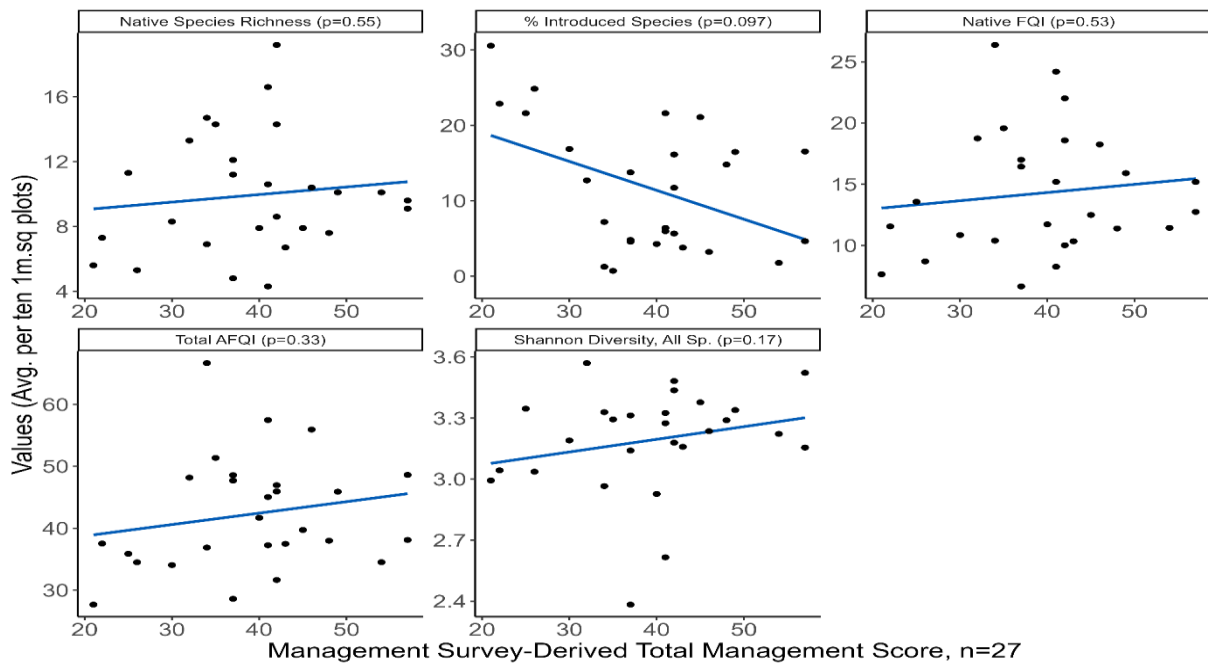


Figure B13. Line graphs depicting the five response variables when compared to the score of management intensity, demonstrating that there are general trends following in the increasing of management, but they are significantly less statistically-relevant than the correlations to staffing scores. Based on existing literature, we would expect to see a correlation between increasing management intensity and increasing floral diversity and quality (Steinauer and Collins 1996; Wagle and Gowda 2018; Bach and Kleiman 2021; Reemts and Belcher 2022). While my results only vaguely capture this trend, what was quite clear in my study is that increasing management intensity can reduce the percentage of invasive plant species and increase gamma diversity in prairies, which is consistent with existing literature (Campbell and Hooymans 2016; Sluis, Bowles, and Jones 2018; Wagle and Gowda 2018; Bach and Kleiman 2021). It would seem that, again, management and labor has a greater impact than other external variables.

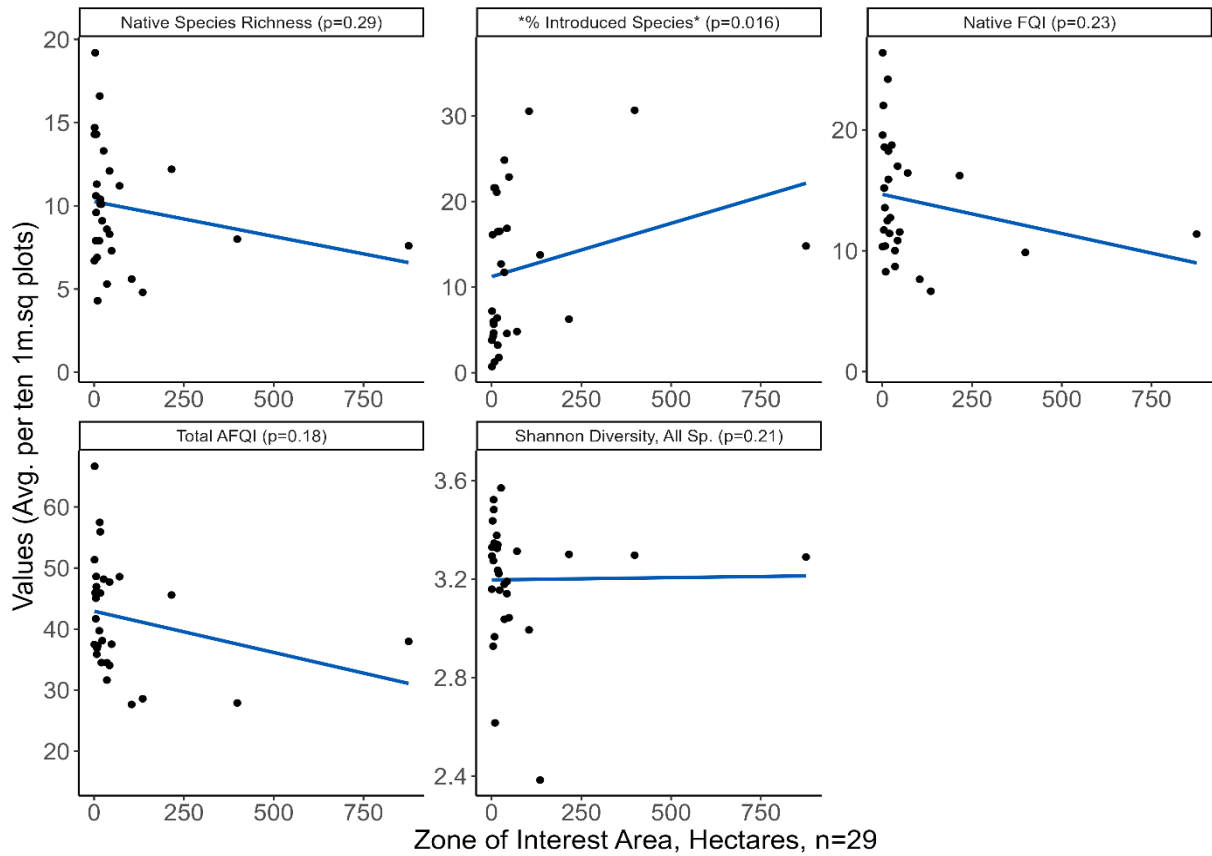


Figure B14. Line graphs depicting the five response variables when compared to hectare size of each zone of interest. There are some correlations and trends, but the only significant one is the increasing percent of invasive species with increasing size. General theory holds that conserved sites away from suburban zones tend to be bigger and older, and bigger and older conserved sites tend to have higher diversity and quality. (Polley, Derner, and Wilsey 2005; Grman, Bassett, and Brudvig 2013; Rowe 2013; Wagle and Gowda 2018). However, this is not always the case in tallgrass prairies (Howe 1994; Kaul and Wilsey 2021). Similarly, this was not the case in my results, as I analysed and found that there were no correlations between the size or age of a prairie and its quality.

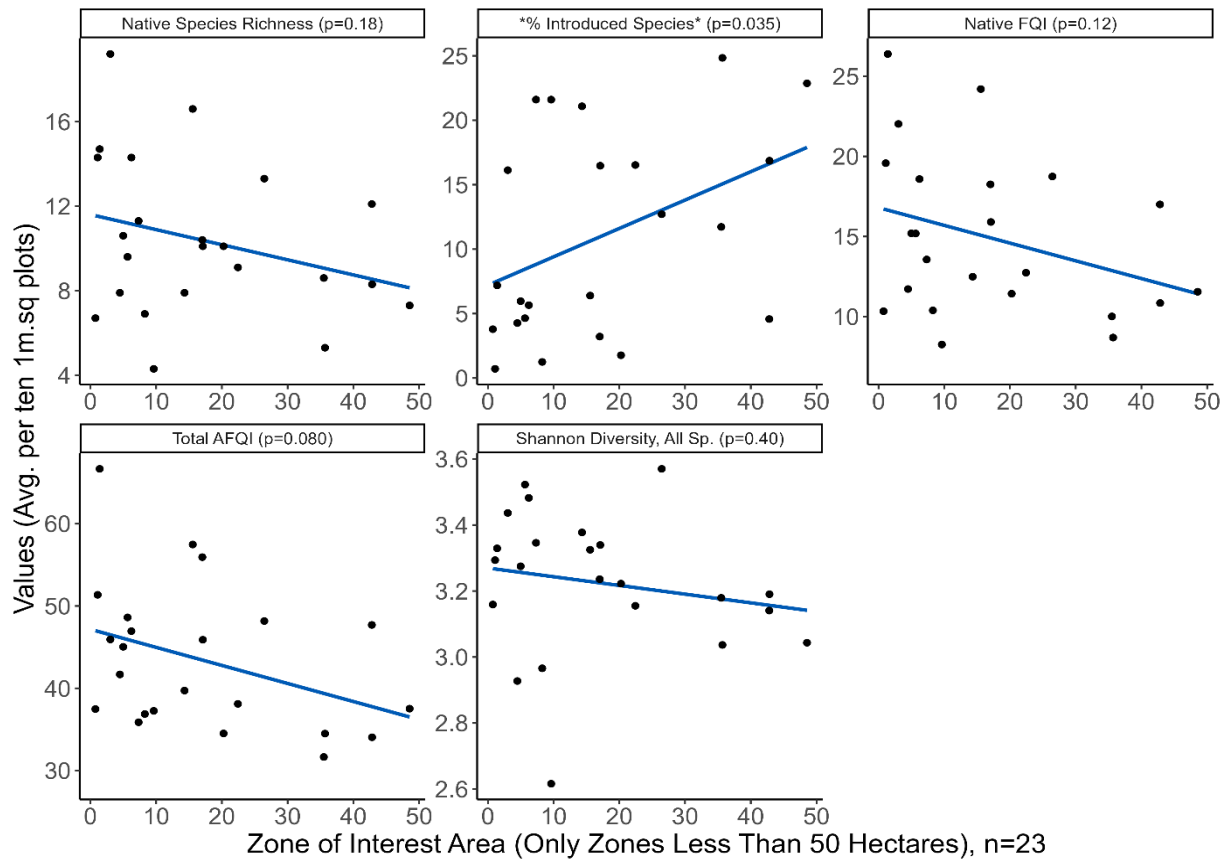


Figure B15. Line graphs depicting the five response variables when compared to hectare size of each zone of interest, though the largest sites have been dropped so this only compares zones less than 50 hectares. There are some correlations and trends, but the only significant one is the increasing percent of invasive species with increasing size, same as when all sizes of zones were compared. As introduced species are undesirable in prairie conservation areas, citing a “quality increase” for introduced species refers to a numeric decrease in the percentage of introduced species. An increase in introduced species would be the inverse, a quality decrease.

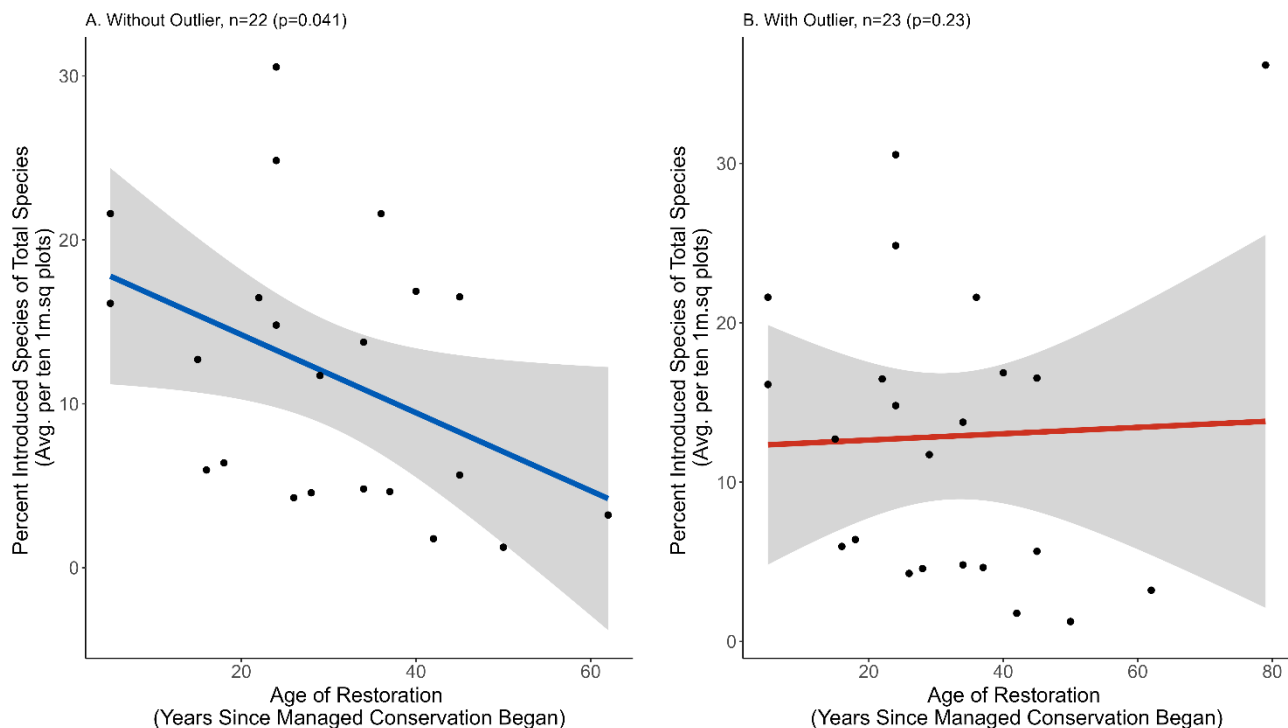


Figure B16. Two line graphs demonstrating the significant changes introduced by the inclusion/exclusion of the outlier zone, Trelease Prairie (zone 3). Figure A on the left is without Trelease, and Figure B on the right is with Trelease (which is the point in the upper right corner). As Trelease's presence was highly weighted and radically changed results in some measurements, but had very little impact in others, I determined that its removal would positively affect statistical analysis. In measurements where it had little impact, its absence is barely felt. In measurements where it had greater impact, its absence allows for more accurate data interpretation.

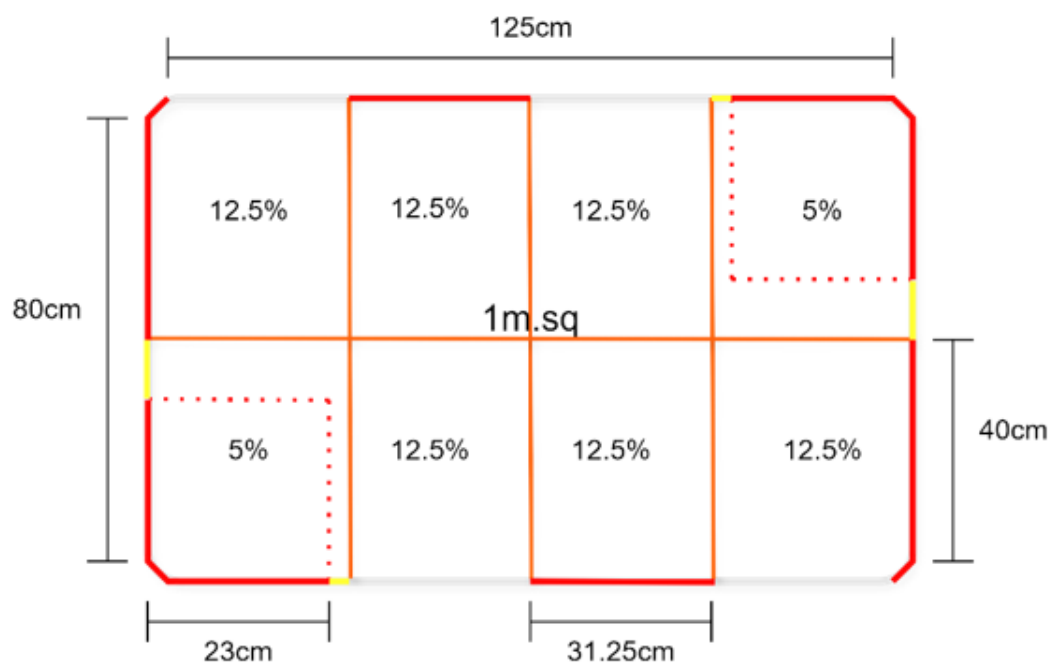


Figure B17. Diagram depicting the exact measurements of my survey quadrat I used at my plots to define the 1 m² sampling area. Although it is not a 1 by 1 m square, its area is still 1 m². Percentages represent the coverage of a subunit within the quadrat and were used for Daubenmire estimation (Higgins et al. 1996). For example, a species that was present in three subunits might have a coverage around 37.5%, giving it a class of 4.

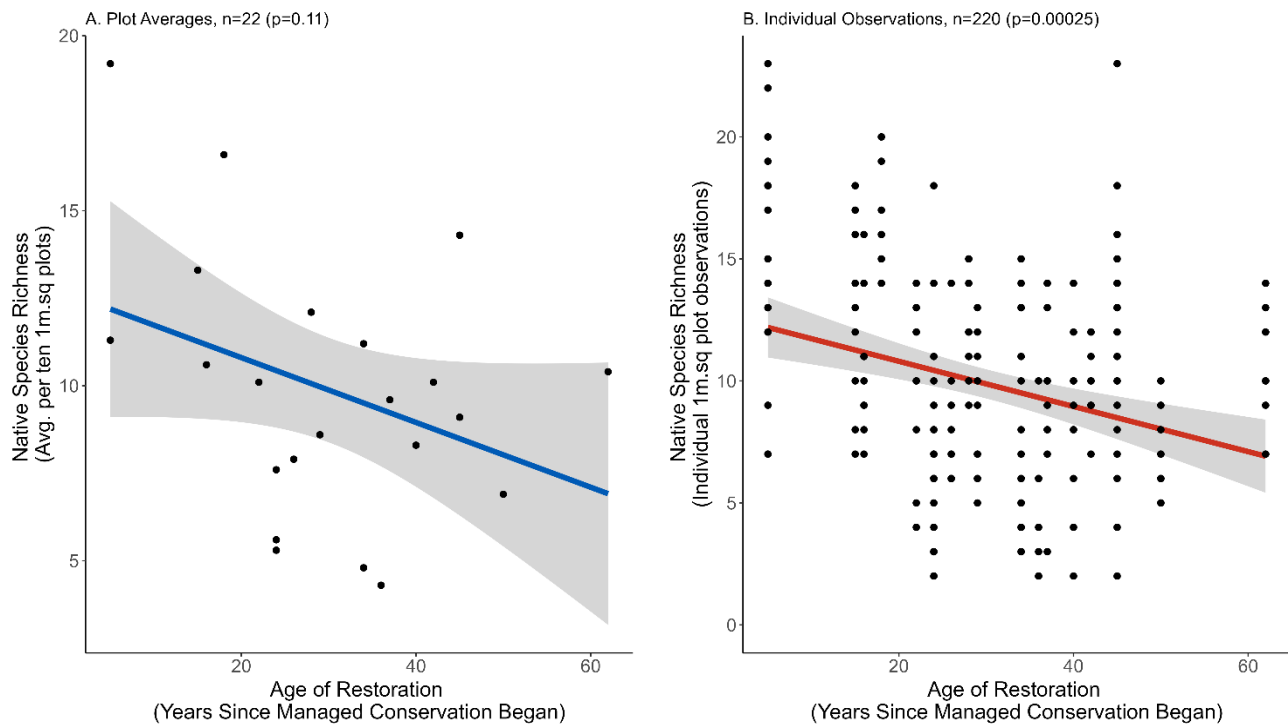


Figure B18. Two line graphs demonstrating the lack of difference in trend and magnitude between zone averages of ten plots (left) and the individual plot observations themselves (right). Because plot observations in each zone all have the same predictor variables, they cannot be effectively analyzed as individual observations. However, there is minimal difference between trends observed by analyzing individual observations versus plot averages, making the distinction meaningless anyway. When running a GLM process on individual plot observations ($n=230$), similar to section 3.4, generating a very large multivariate model suggested that organizational operating budget and grouped age of restoration were additionally quite impactful, though I did not analyse this within a smaller model. While this difference is interesting, I determined that it was not generally meaningful to analyse individual plot observations. Individual observations lack the statistical power to be analysed on their own against zone or site-level predictor variables. Plots are a singular square meter of prairie and should be aggregated to get a more comprehensive picture of the whole prairie, especially as zone- or site-wide predictor variables represent an entire prairie. That being said, what analyses I did perform suggested that the shape and trends of the results are unchanged by analysing individual observations instead of plot averages, as demonstrated above. As land cover and management values were assigned at the site and/or zone level, each sample plot within a zone had exactly the same predictor variables, meaning that intra-zone variability was zero and the group boundary produced a singularity, further precluding the use of mixed model analysis (Zuur et al. 2009).

Appendix C: Sample Management Survey

The following is a text copy of the management survey that was sent to land managers responsible for the study sites sampled within this project. The survey was originally sent to land managers as both a word document and as a Google Forms survey, though content was identical in both. Also attached is some of the explanatory material given to respondents in order to preemptive answer any questions.

Question Guide to the Primary Digital Questionnaire

For Research and Data Collection on Northern Illinois Tallgrass Prairie Biodiversity And Community Structure – Question Guide

Survey Research Question: What are the distribution and value of independent variables affecting tallgrass prairie biodiversity across thirty tallgrass prairie conservation sites in northeastern Illinois, as reported by relevant land managers? How do these variables (management technique and intensity, age of prairie restoration, and organizational budget and resource expenditure) differ between sites? This data will be used as part of a broader research subject (see explanatory document).

Specific Purpose of Primary Digital Questionnaire

My name is Andy Sima, and I am a master's student at Stockholm University studying Landscape Ecology at the Department of Physical Geography (<https://www.su.se/departments-of-physical-geography/>). I am originally from Chicago, and received my bachelor's degree in environmental sustainability from the University of Illinois at Urbana-Champaign.

Over August and September of 2024, I have sampled floral biodiversity at a number of northeastern Illinois tallgrass prairie sites, which may or may not include your own, permit status-dependent. I identified every plant species present and visually estimated coverage of species within ten 1-sq.meter plots. The plots were distributed randomly within a “zone of interest” (see attached map), or a mostly-contiguous patch of prairie. To understand age and management of these sites, I would like to ask you a few questions.

I am requesting that you fill out this Google Form questionnaire. There is also a word document copy contained within the attached explanatory document. It is my estimation that the questionnaire will take between 10-30 minutes per site. I may be requesting you to fill out this questionnaire for several zones of interest. Please refer to the map(s) included with the original email for specific zones of interest/sites and fill out one questionnaire each per zone as soon as possible, but preferably no later than December 16th, 2024. Please feel free to send this to other relevant land managers for the zone.

For more information about this questionnaire, study-specific methodology and terminology, concerns you may have about how to answer these questions, your rights as a respondent, and more, **please refer to the attached explanatory document**. If you wish to provide more information beyond the scope of this questionnaire via an **optional secondary questionnaire**, or want to schedule a phone/zoom interview to further explain your answers, please email me directly.

Thank you for your time.

Primary Digital Questionnaire – Questions Crucial to This Study

Interviewee Identification Subsection

Your rights as a questionnaire respondent: As a part of answering this questionnaire, your responses will remain anonymous and de-identified in the final report. You also have the further right to request your answers be redacted or removed entirely at a later date (prior to project completion in May 2025), to refuse to answer any question, and to cease the questionnaire at any time with no negative consequences. Please see the attached explanatory document for more information regarding your rights and questionnaire ethics.

Furthermore, by submitting answers via a Google Form, submitted information does enter into Google's data storage system and is, therefore, subject to their privacy and data use policies. **If you prefer to avoid this, please answer questions in the attached word document instead and email it back to me directly.**

Furthermore, some questions regarding organizational budget and expenses may, in some unlikely cases, pose confidentiality risks, according to your organization's policies. Please keep in mind if you are under a non-disclosure agreement or are restricted in discussing your work in any way, and **feel free to skip any questions which may prove to be objectionable.**

1. **Interviewee name:** *your answers will be anonymized in the final report. long-form answer*
2. **Employing Organization:** *long-form answer*
3. **Contact Information:** *what's the best way to reach you if I need to follow-up? For research purposes, email is preferable. This question is skippable. long-form answer*
4. **What is your role at this site? What is your official position?** *long-form answer*
5. **How long have you worked at your current position in this site?:** *Preferably, you have worked at this site long enough to develop a working knowledge of this prairie's quality and history, or otherwise have access to relevant internal documents. If you do not feel you have adequate site knowledge regarding management/restoration to answer these questions, please send this questionnaire to a colleague who does. Even if you **do** have proper site knowledge, I would appreciate it if you would still send this questionnaire to other colleagues at your site in addition to answering it yourself. long-form answer*
6. **I, the above stated, grant Andy Sima the permission, in good faith, to record, reproduce, analyze, disseminate, and/or publish in full, anonymized, the answers contained within this document for the purpose of scientific study and research, including the scientific publication of results generated from this information. My answers in this document will be accurate to the best of knowledge.: Yes or no**

Site Identification Subsection

Questions designed to help clarify the status of the site (blue area on the map, attached in original email) and zone of interest (yellow area on the map, attached in original email) you are filling out this questionnaire for. Please fill out a questionnaire for **each requested zone of interest**.

1. **I am answering this questionnaire for site number...** *(the entire contiguous protected zone; check the attached map for the site number and names; site is in blue)*
drop-down list of sites 1-30
2. **I am answering this questionnaire for zone of interest number...** *(specific prairie patch and study unit; see explanatory note for more information, and please answer questionnaire questions for only one zone of interest; check the attached map for the site number and names; zone of interest is in yellow. Site name and zone of interest are not always the same.)* *drop-down list of sites 1-30*
3. **What is your site's official name?:** *long-form answer*
4. **What is the name of the organization that owns this site?:** *long-form answer*
5. **Is the organization that owns this site public or private sector?:** *public sector, private sector, don't know, other*
6. **What is the name of the organization that manages this site?** *if separate entity than ownership organization, i.e. contract work. If not separate, please leave this question blank.* *long-form answer*
7. **Does your site (the entire contiguous protected space) or zone of interest (specific prairie patch denoted in the attached map from the email; see separate explanatory note for further definitions) have any special legal designations?:** *e.g. Illinois Nature Preserve Area, National Wildlife Refuge, etc.,* *long-form answer*

Management Practices Subsection

All questions within this segment apply only to the designated **zone of interest** *(the yellow patch in the attached map; see explanatory document for more information).*

1. Does this zone of interest contain remnant prairie (*never plowed, developed, or significantly disrupted in its growth patterns*)?: *yes or no*
2. Does this zone of interest contain restored prairie (*original prairie entirely plowed over or removed, but later replanted*)?: *yes or no*
3. On a scale of 1 to 10, with 1 being “entirely restoration” and 10 being “entirely remnant,” would you consider the prairie within this zone of interest predominantly remnant or restoration?: *1-10 scale*
4. If this zone of interest contains restoration, what year did restoration commence? *to the specific year, if possible. Beginning of restoration is here defined as when the earliest restored seedbank was established, i.e. the first time prairie seeds were re-introduced. If remnant, please state “remnant.” long-form answer*
5. If this zone of interest contains restoration, what is this site’s previous land use history? *Agriculture, exceptionally destructive rangeland/grazing, industrial development, residential development, unknown, other*
6. What are the primary management techniques for this zone of interest? Mark all that apply: *checkboxes; burning, mowing or haying, livestock grazing, seeding, pesticides/herbicides, hand weeding, don’t know, other (please explain)*
7. On average, what is the frequency of this zone of interest’s management practices?

	Do not use this technique	Don’t Know	Several times a year	Once a year	Every 2-4 years	Every 5-10 years	Greater than 10 years
Burning							
Mowing or Haying							
Livestock Grazing							
Seeding							
Pesticides or Herbicides							
Hand Weeding							
Other:							

8. If you answered “other” to the above question, or wish to add more, please explain.: *Please note that this and future “other/add more” questions can be skipped without detriment to the rest of the information entered in this survey. long-form*

answer

9. **What percentage of this zone of interest's total area is typically disturbed during an application of any given technique?** *I.e. given the zone of interest, if you are choosing to manage it with fire, how much do you burn at a time? Half of it? When you mow, how much of the zone is mowed at once? If you use pesticides/herbicides, is it a wide-scale application or spot-treatment? Etc.*

	Do not use this technique	Don't Know	Spot- Treat Problem Areas (less than 5% of area)	Between 5% and 25%	Between 25% and 50%	Between 50% and 75%	Greater than 75% At a Time
Burning							
Mowing or Haying							
Livestock Grazing							
Seeding							
Pesticides or Herbicides							
Hand Weeding							
Other							

10. If you answered "other" to the above question, or wish to add more, please explain.: *long-form answer*

11. **How long ago did your site begin application of this technique to this zone of interest?**

	Do not use this technique	Don't know	Less than 5 years ago	Between 5 and 10 years ago	Between 10 and 20 years ago	Between 20 and 50 years ago	Greater than 50 years ago
Burning							
Mowing or Haying							
Livestock Grazing							

Seeding							
Pesticides or Herbicides							
Hand Weeding							
Other							

12. If you answered “other” to the above question, or wish to add more, please explain.: *long-form answer*
13. On a scale of 1 to 10, with 1 being very low management and 10 being very high management, what is your personal estimation of the intensity of this zone of interest’s management?: *how often do people work on this zone of interest? Is it highly-managed or allowed to run its course? This is a purposefully subjective question. 1-10 scale*
14. On a scale of 1 to 10, with 1 being very unsuccessful and 10 being very successful, how successful has restoration/management been for this zone of interest since you began your job at this site?: *have your restoration and management goals been met? 1-10 scale*
15. Has this zone of interest experienced any major non-management disturbances within the last five years? *i.e. flood, extreme wildfire, tornado, exceptional vandalism, etc. long-form answer*
16. Are there any other relevant management regimes, techniques, or specifics about your site and this zone of interest?: *i.e. if this zone of interest has a very unusual or intense management regime compared to other sites, if especially large-scale or unusual machinery is utilized here, or anything that didn’t fit previous questions, etc. This question can be skipped. long-form answer*

Socioeconomic Information Subsection

Questions here refer to the site as a whole, unless otherwise stated. Some information requested from these questions may be classified or sensitive; please note if so. ’

1. What is the number of permanently-employed/full-time staff with direct land management responsibilities attending to this site?: *i.e. how many permanent paid staff have worked on this site in the past year? Please include staff who are not directly assigned to this site, but do manage this site as part of their job (i.e. rotating staff between sites); this may include maintenance staff, but do not include administrative or interpretive staff unless they are directly involved in preservation. No permanent staff; this site is largely unmanaged, Between 1 and 3 permanent staff,*

between 3 and 5, between 5 and 7, between 7 and 10, between 10 and 15, between 15 and 20, between 20 and 30, between 30 and 50, more than 50 permanent staff, don't know, other

2. **What is the number of seasonal staff with direct land management responsibilities attending to this site?:** *i.e. paid personnel who only work for the summer, but are not volunteers. May include staff that rotate between sites but are directly involved with this site. No seasonal staff, Between 1 and 3 seasonal staff, between 3 and 5, between 5 and 7, between 7 and 10, between 10 and 15, between 15 and 20, between 20 and 30, between 30 and 50, more than 50 seasonal staff, don't know, other*
3. **What is the number of volunteers with direct land management responsibilities attending to this site?:** *i.e. unpaid individuals who work regularly, perhaps weekly or monthly, or perhaps only once a year, on prairie management and restoration at this site, either through special volunteer relationships or via designated volunteer events. No volunteers, Between 1 and 3 volunteers, between 3 and 5, between 5 and 7, between 7 and 10, between 10 and 15, between 15 and 20, between 20 and 30, between 30 and 50, more than 50 volunteers, don't know, other*
4. **Are there any other relevant sources of man-power?:** *contractor companies, etc. long-form answer*
5. **In a given year how often do you, personally, work on this site?** *Please only consider days that are spent in active management, i.e. out in the field. I do not typically work at this site, I rarely work at this site (a few times a year), I work at this site sometimes, I work at this site monthly, I work at this site weekly, I work at this site daily, don't know, other*
6. **What is the current estimated financial expenditure for only this site as a whole?:** *What is the conservation and management budget for this site (including the zone of interest)? How much money do you estimate that your organization is spending on this site, including staff, resources, construction, maintenance, etc.? Less than \$100,000, Between \$100,000 and \$500,000, Between \$500,000 and \$1 million, Between \$1 and \$2.5 Million, between \$2.5 and \$5 million, between \$5 and \$10 Million, between \$10 and \$15 Million, between \$15 million and \$20 million, between \$20 and \$30 million, between \$30 and \$40 million, Between \$40 and \$50 million, between \$50 and \$75 million, between \$75 and \$100 million, greater than \$100 million, don't know, other*
7. **What are the primary sources of funding for this site?** *Checkboxes; Tax base/local and county government allocation, Tax base/state government allocation, tax base/federal government allocation, public grants (i.e. NSF), private grants, private*

donations, sales or site revenue generation, don't know, other

8. **On a scale of 1 to 10 with 1 being “very low damage” and 10 being “very high damage,” how would you categorize damage to this site’s ecosystems by the public’s usage of this site? Please put zero if this site is not open to the public. 0-10 scale**
9. **What is the direct overhead organizational entity for this site?: i.e., what is the organization responsible for the non-land management duties of running this site? It may be the same organization as you work for, or it may not. Duties would include logistical, administrative, etc. long-form answer**
10. **What is the estimated number of other conservation sites managed by this overhead organizational entity, beyond your site?: Only this site, 1 to 2 other sites, 3 to 5 other sites, 5 to 10 other sites, 10 to 15 other sites, 15 to 25 other sites, 25 to 40 other sites, 40 to 60 other sites, 60 to 80 other sites, 80 to 100 other sites, greater than 100 other sites, don't know, other**
11. **What is the estimated total operating budget of your overhead organization?: how much money does this organization spend in a given year, on all expenses, not just conservation and land management? Less than \$100,000, Between \$100,000 and \$500,000, Between \$500,000 and \$1 million, Between \$1 and \$2.5 Million, between \$2.5 and \$5 million, between \$5 and \$10 Million, between \$10 and \$15 Million, between \$15 million and \$20 million, between \$20 and \$30 million, between \$30 and \$40 million, Between \$40 and \$50 million, between \$50 and \$75 million, between \$75 and \$100 million, between \$100 and \$150 million, between \$150 and \$250 million, between \$250 and \$300 million, between \$300 and \$400 million, between \$400 and \$500 million, between \$500 million and \$1 billion, greater than \$1 billion, don't know, other**
12. **What are the primary sources of funding for your overhead organization? Checkboxes; Tax base/local and county government allocation, Tax base/state government allocation, tax base/federal government allocation, public grants (i.e. NSF), private grants, private donations, sales or site revenue generation, don't know, other**
13. **Do you have any further comments on this site’s organizational structure and socioeconomic makeup?: a space for anything you feel is relevant but didn't fit other questions. This question can be skipped. long-form answer**
14. **Do you have any other comments about the zone of interest, your site in general, or anything else you wish to add that was not in a question but may be relevant? This question can be skipped.**

End of Questionnaire

Thank you for your time on this questionnaire and for your efforts in maintaining and restoring a special ecosystem. At the conclusion of this project, all respondents will receive a complete set of anonymized data I collected as part of this project. If you have any further questions, want to answer questions for the secondary questionnaire, or wish to schedule a phone/zoom interview, please feel free to reach out.

1. Did you choose to skip any questions due to the risk of answers containing potentially sensitive information? Did you feel that any of the questions posed herein required you to breach any non-disclosure agreements or restrictions to which you may be obligated? Please elaborate as necessary in the “other” category. *Yes or no*

Clarification of Study-Specific Terms and Methodologies

Clarifying Terms – “Site” refers to the overall contiguous area and land ownership unit that the prairie is contained within, such as the entire Forest Preserve or State Park area.

“Zone of Interest” refers to the specific segment of prairie within the site, as delineated by the attached zone of interest map. Some sites may have two separate zones of interest, for which two separate questionnaires should be completed.

“Remnant” here defines prairies as remnants if they have never been plowed, converted into agriculture, or had an interruption in the continuous existence of grassland.

“Restoration” refers to all other prairies, i.e. those that have been restored to grassland since plowing/conversion to agriculture. If you feel your site does not fit one of these categories, please explain further.

On Zones of Interest - For this study, I have semi-randomly (via geographic stratification and publicly available data) selected 30 prairies sites within northern Illinois, and within some these 30 prairies sites I have outlined specific geographic zones of interest. This questionnaire is accompanied by a digital map for your site outlining the extent of this prairie zone in question, as the zone of interest may not representative of the “entire” prairie for which you/your organization are responsible for. Questions may be either site- or zone-dependent. The boundaries of your site, as detailed in the zone of interest map, are gathered from the Prairie State Conservation Coalition’s GIS Database of preserves spaces, (<https://pscc.maps.arcgis.com/apps/webappviewer/index.html?id=fb8dbaacc7b04c7fa12e850b9e7141ab>) by David Holman.

On Collaboration in Answering Questions –The answering of this questionnaire can absolutely be a collaborative effort between several staff members or individuals within a department, if desired. It is my intent to gather the input of as many relevant individuals as possible for each site, but if you want to collaborate on one form for you site, please include

information for all contributing individuals. **However, for sites with multiple zones of interest, please limit answers only one zone of interest per questionnaire.**

On Levels of Detail for Answers – For ease of use, most question answer options have been categorized into distinct chunks. However, for long-answer questions, please answer questions with the level of detail you feel is relevant to the question, as you might in an in-person interview. This may just be a yes or no, or a list of management techniques. Or this may be several sentences explaining why the answer for your specific site is complicated, or a few sentences of your personal opinion on a question.

On Sites or Individuals With Two or More Zones of Interest – Some of these sites have two zones of interest (which will be demonstrated in the attached imagery). Fill out one questionnaire per zone of interest; if your site contains two zones, please fill out two questionnaires. Some site-wide questions may have significant overlap in their answers between zones of interest; feel free to copy and paste applicable answers between questionnaires. Furthermore, you may be responsible for two or more of the zones of interest I have selected, even beyond one single site. **Please fill out a new questionnaire for every zone of interest you have responsibility for.**

On Multi-Use Zones of Interest – If a designated zone of interest contains more than one type of prairie or restoration, try to be as specific as possible with delineations among prairie age, management type, etc. When in doubt, summarize answer to the most general area of similar maintenance/shared management unit. For example, if a working unit contains two restorations, aged five years apart, but they are now treated as one management unit, include available information on separate restoration age and then summarize management information to the scale of the working unit.

On Unknown Or Classified Information – If there are no records answering some of the questions posed here (i.e. no records of original source of the seed bank), put “no records” or something similar. If information is considered sensitive, or answers to a question would be considered against your organization’s policy, please put “sensitive info” or “choose not to answer.” If you are comfortable with sharing information on this questionnaire but *not* comfortable with it being distributed post-completion, please make note and clarify what information ought to be redacted before sharing this data with other participants.

On Using this Word Document – The primary questionnaire word document is identical to the accompanying Google Form with questions. Please answer questions directly into the Google Form and please return one set of answers per respondent, per requested site. If you wish to supply information about your site via the optional secondary questionnaire, answers can be entered into the word document directly and returned as .docx or .pdf file, but please enter answers in a manner distinct from the body of questions. This may either be an entire separate document, or using a red font to answer questions, etc.

On Your Rights as a Questionnaire Respondent – As a part of answering this questionnaire, you have a right to remain anonymous, to request your answers be redacted (or removed entirely) at a later date (prior to project completion in May 2025), refuse to answer any question, or cease the questionnaire at any time with no negative consequences. You also have the right to request that any identifying factors relating you to this data be destroyed prior to full publication. As a student thesis project, this questionnaire is not subject to Stockholm University Ethics Review but is subject to all applicable Swedish laws regarding

the collection of personal data and information, and is subject to Stockholm University student studies ethical conduct. For more information regarding ethical conduct of research interviews through Stockholm University, please see this link to Stockholm University's research ethics page (<https://www.su.se/english/about-the-university/university-facts/research-integrity-and-ethics>) and this example of a master's thesis ethics requirement document.

(https://www.su.se/polopoly_fs/1.606196.1648633248!/menu/standard/file/Declaration%20of%20ethics%20for%20student%27s%20thesis%20work.pdf)

You also have the right to know that this data is being collected for the purpose of scientific analysis, research, and dissemination in the specific field of grassland conservation. Data will be aggregated, linked to relevant sites, and analyzed with further datasets beyond that which is supplied by respondents.

Furthermore, by submitting answers via a Google Form, submitted information does enter into Google's data storage system and is, therefore, subject to their privacy and data use policies. If you prefer to avoid this, please answer questions in this word document instead and email it back to me directly.

On The Expected Future of this Research – Upon completion of this master's thesis, the data and results of this project, along with the completed report, will be logged in Stockholm University's online archival system. It is my intention that this data and report be made freely available to the wider scientific community for future usage. Depending on the relevance and usefulness of results, this report may or may not be published academically in a scientific journal. Furthermore, as a creative writer as well as an ecologist, I intend to write a memoir or literary essay detailing my experiences in the prairies and my experiences working on this thesis project.

Here is a link to this explanatory document, hosted online. It exists for reference and to add context: https://drive.google.com/file/d/1vvT4AlupaBMJ8hCwCc37KFQg_-0C_RS7/view?usp=sharing

Appendix D: Full Species List

Total Number of Unique Species (not including unidentified): 232 species (195 native, 37 introduced)

Top Ten Highest Frequency/Coverage Native Species Across All Samples: *Solidago altissima* (tall goldenrod), *Sorghastrum nutans* (Indian grass), *Andropogon gerardii* (big bluestem), *Monarda fistulosa* (wild bergamot), *Ratibida pinnata* (gray-headed coneflower),

Zizia aurea (golden alexander), *Coreopsis tripteris* (tall coreopsis), *Euthamia graminifolia* (grass-leaved goldenrod), *Oligoneuron rigidum* (stiff leaved goldenrod), *Cornus racemosa* (gray dogwood)

Top Ten Highest Frequency/Coverage All Species Across All Samples: *Solidago altissima* (tall goldenrod), *Sorghastrum nutans* (Indian grass), *Andropogon gerardii* (big bluestem), *Monarda fistulosa* (wild bergamot), *Bromus inermis* (smooth brome), *Ratibida pinnata* (gray-headed coneflower), *Zizia aurea* (golden alexander), *Coreopsis tripteris* (tall coreopsis), *Euthamia graminifolia* (grass-leaved goldenrod), *Oligoneuron rigidum* (stiff leaved goldenrod)

Scientific names per Mohlenbrock's *Vascular Flora of Illinois* (2014). C values and nativity status per Wilhelm and Rericha's *Flora of the Chicago Region* (2017). Some species are native to the North American continent but are defined as introduced to the tallgrass prairie ecosystem by Wilhelm and Rericha. These are denoted by "Introduced (N. America)."

Scientific Name	Common Name	Nativity Status	C Value
<i>Achillea millefolium</i>	common yarrow	Native	0
<i>Ageratina altissima</i>	white snakeroot	Native	3
<i>Agrimonia gryposepala</i>	tall hairy agrimony	Native	2
<i>Agrimonia parviflora</i>	swamp agrimony	Native	4
<i>Agrostis gigantea</i>	redtop bentgrass	Introduced	0
<i>Allium cernuum</i>	nodding onion	Native	7
<i>Ambrosia artemisiifolia</i>	common ragweed	Native	0
<i>Ambrosia trifida</i>	giant ragweed	Native	0
<i>Amorpha canescens</i>	leadplant	Native	10
<i>Amphicarpaea bracteata</i>	American hog-peanut	Native	5
<i>Andropogon gerardii</i>	big bluestem	Native	5
<i>Anemone canadensis</i>	meadow anemone	Native	4
<i>Anemone cylindrica</i>	cylindrical thimbleweed	Native	8
<i>Anemone virginiana</i>	tall thimbleweed	Native	5
<i>Apios americana</i>	American groundnut	Native	5
<i>Apocynum cannabinum</i>	common dogbane	Native	6
<i>Arnoglossum plantagineum</i>	prairie Indian plantain	Native	10
<i>Asclepias syriaca</i>	common milkweed	Native	0
<i>Asclepias tuberosa</i>	butterfly milkweed	Native	8
<i>Asclepias verticillata</i>	whorled milkweed	Native	1
<i>Astragalus canadensis</i>	Canadian milkvetch	Native	8
<i>Baptisia alba</i>	white wild indigo	Native	8
<i>Baptisia australis</i>	blue wild indigo	Introduced (N. America)	0
<i>Baptisia bracteata</i>	cream wild indigo	Native	10
<i>Blephilia hirsuta</i>	hairy wood mint	Native	5
<i>Botrypus virginianus</i>	rattlesnake fern	Native	5
<i>Bouteloua curtipendula</i>	sideoats grama	Native	8
<i>Brickellia eupatorioides</i>	false boneset	Native	6
<i>Bromus inermis</i>	smooth brome	Introduced	0
<i>Calamagrostis canadensis</i>	Canada reed grass	Native	6
<i>Calystegia sepium</i>	hedge bindweed	Native	1
<i>Carex cristatella</i>	crested sedge	Native	4
<i>Carex pellita</i>	wooly sedge	Native	4

<i>Carex scoparia</i>	pointed broom sedge	Native	5
<i>Carex vulpinoidea</i>	fox sedge	Native	2
<i>Ceanothus americanus</i>	New Jersey tea	Native	8
<i>Celastrus scandens</i>	American bittersweet	Native	4
<i>Chaerophyllum procumbens</i>	wild chervil	Native	5
<i>Chamaecrista fasciculata</i>	partridge pea	Native	4
<i>Cirsium altissimum</i>	tall thistle	Native	8
<i>Cirsium arvense</i>	creeping thistle	Introduced	0
<i>Cirsium discolor</i>	field thistle	Native	3
<i>Comandra umbellata</i>	bastard toadflax	Native	9
<i>Coreopsis palmata</i>	prairie coreopsis	Native	10
<i>Coreopsis tripteris</i>	tall coreopsis	Native	5
<i>Cornus racemosa</i>	gray dogwood	Native	1
<i>Cornus sericea</i>	red osier dogwood	Native	5
<i>Dalea candida</i>	white prairie clover	Native	10
<i>Dalea purpurea</i>	purple prairie clover	Native	9
<i>Daucus carota</i>	wild carrot	Introduced	0
<i>Desmanthus illinoensis</i>	illinois bundleflower	Native	3
<i>Desmodium canadense</i>	showy tick-trefoil	Native	4
<i>Desmodium illinoense</i>	illinois tick-trefoil	Native	9
<i>Dichanthelium oligosanthos</i>	heller's rosette grass	Native	9
<i>Dichanthelium scribnerianum</i>	scribner's panic grass	Native	4
<i>Drymocallis arguta</i>	tall cinquefoil	Native	10
<i>Echinacea pallida</i>	pale purple coneflower	Native	10
<i>Echinacea purpurea</i>	purple coneflower	Native	10
<i>Elymus canadensis</i>	Canada wild rye	Native	4
<i>Elymus virginicus</i>	Virginia wild rye	Native	3
<i>Epilobium coloratum</i>	cinnamon willowherb	Native	3
<i>Epilobium leptophyllum</i>	linear-leaf willowherb	Native	10
<i>Equisetum arvense</i>	common horsetail	Native	0
<i>Erigeron annuus</i>	annual fleabane	Native	0
<i>Eryngium yuccifolium</i>	rattlesnake master	Native	9
<i>Eupatorium altissimum</i>	tall boneset	Native	0
<i>Eupatorium serotinum</i>	late boneset	Native	0
<i>Euphorbia corollata</i>	flowering spurge	Native	4
<i>Euphorbia cyparissias</i>	cypress spurge	Introduced	0
<i>Euthamia graminifolia</i>	grass-leaved goldenrod	Native	4
<i>Fragaria virginiana</i>	wild strawberry	Native	0
<i>Frangula alnus</i>	alder buckthorn	Introduced	0
<i>Fraxinus pennsylvanica</i>	green ash	Native	4
<i>Galium mollugo</i>	white bedstraw	Introduced	0
<i>Galium obtusum</i>	bluntleaf bedstraw	Native	5
<i>Gaura biennis</i>	biennial beeblossom	Introduced (N. America)	0
<i>Gentiana alba</i>	pale gentian	Native	9
<i>Gentiana andrewsii</i>	closed bottle gentian	Native	9
<i>Gentiana puberulenta</i>	downy gentian	Native	10
<i>Gentianella quinquefolia</i>	agueweed	Native	8

<i>Geum aleppicum</i>	yellow avens	Native	3
<i>Geum canadense</i>	white avens	Native	1
<i>Geum laciniatum</i>	rough avens	Native	3
<i>Hackelia virginiana</i>	Virginia stickseed	Native	1
<i>Helenium autumnale</i>	common sneezeweed	Native	5
<i>Helianthus decapetalus</i>	pale sunflower	Native	8
<i>Helianthus divaricatus</i>	woodland sunflower	Native	5
<i>Helianthus grosseserratus</i>	sawtooth sunflower	Native	4
<i>Helianthus mollis</i>	ashy sunflower	Native	9
<i>Helianthus pauciflorus</i>	stiff sunflower	Native	9
<i>Helianthus tuberosus</i>	jerusalem artichoke	Native	3
<i>Heliopsis helianthoides</i>	false sunflower	Native	7
<i>Heuchera richardsonii</i>	prairie alumroot	Native	10
<i>Hypericum ascyron</i>	great St. John's wort	Native	10
<i>Hypericum perforatum</i>	common St. John's wort	Introduced	0
<i>Hypericum punctatum</i>	spotted St. John's wort	Native	4
<i>Hypericum sphaerocarpum</i>	roundfruit St. John's wort	Native	7
<i>Isanthus brachiatus</i>	false pennyroyal	Native	7
<i>Juglans nigra</i>	eastern black walnut	Native	3
<i>Juncus greenii</i>	greenie's rush	Native	8
<i>Juncus tenuis</i>	slender path rush	Native	0
<i>Juncus torreyi</i>	torrey's rush	Native	2
<i>Lactuca biennis</i>	tall blue lettuce	Native	5
<i>Lespedeza capitata</i>	round-headed bush clover	Native	4
<i>Lespedeza cuneata</i>	Chinese bushclover	Introduced	0
<i>Leucanthemum vulgare</i>	oxeye daisy	Introduced	0
<i>Liatris aspera</i>	rough blazing star	Native	8
<i>Liatris pycnostachya</i>	prairie blazing star	Native	8
<i>Liatris spicata</i>	marsh blazing star	Native	7
<i>Lithospermum canescens</i>	hoary puccoon	Native	8
<i>Lobelia siphilitica</i>	great blue lobelia	Native	4
<i>Lonicera maackii</i>	Amur honeysuckle	Introduced	0
<i>Lotus corniculatus</i>	common bird's-foot trefoil	Introduced	0
<i>Lupinus polyphyllus</i>	large leaved lupine	Introduced (N. America)	0
<i>Lycopus americanus</i>	American bugleweed	Native	4
<i>Lycopus uniflorus</i>	northern bugleweed	Native	4
<i>Lysimachia lanceolata</i>	lanceleaf loosestrife	Native	9
<i>Lythrum alatum</i>	winged loosestrife	Native	7
<i>Medicago lupulina</i>	black medick	Introduced	0
<i>Melilotus alba</i>	white sweetclover	Introduced	0
<i>Melilotus officinalis</i>	yellow sweetclover	Introduced	0
<i>Monarda fistulosa</i>	wild bergamot	Native	4
<i>Monarda punctata</i>	spotted horse mint	Native	4
<i>Morus alba</i>	white mulberry	Introduced	0
<i>Nabalus asper</i>	rough rattlesnake root	Native	10
<i>Napaea dioica</i>	glade mallow	Native	4
<i>Nepeta cataria</i>	catnip	Introduced	0

<i>Oenothera biennis</i>	common evening primrose	Native	0
<i>Oligoneuron ohioense</i>	Ohio goldenrod	Native	10
<i>Oligoneuron riddellii</i>	riddell's goldenrod	Native	8
<i>Oligoneuron rigidum</i>	stiff leaved goldenrod	Native	3
<i>Oxalis stricta</i>	slender yellow woodsorrel	Native	0
<i>Oxypolis rigidior</i>	stiff cowbane	Native	8
<i>Packera paupercula</i>	balsam ragwort	Native	4
<i>Panicum capillare</i>	witch grass	Native	0
<i>Panicum virgatum</i>	switchgrass	Native	3
<i>Parthenium integrifolium</i>	wild quinine	Native	8
<i>Parthenocissus inserta</i>	thicket creeper	Native	0
<i>Pastinaca sativa</i>	wild parsnip	Introduced	0
<i>Pedicularis canadensis</i>	wood betony	Native	9
<i>Penstemon digitalis</i>	foxglove beardtongue	Native	4
<i>Penthorum sedoides</i>	ditch stonecrop	Native	4
<i>Perideridia americana</i>	thicket parsley	Native	8
<i>Phalaris arundinacea</i>	reed canary grass	Introduced	0
<i>Phleum pratense</i>	Timothy grass	Introduced	0
<i>Phlox glaberrima</i>	smooth phlox	Native	9
<i>Phlox paniculata</i>	fall phlox	Native	1
<i>Phlox pilosa</i>	prairie phlox	Native	8
<i>Phryma leptostachya</i>	American lopseed	Native	6
<i>Physostegia virginiana</i>	obedient plant	Native	4
<i>Plantago major</i>	broadleaf plantain	Introduced	0
<i>Populus tremuloides</i>	trembling aspen	Native	3
<i>Potentilla recta</i>	sulphur cinquefoil	Introduced	0
<i>Potentilla simplex</i>	common cinquefoil	Native	3
<i>Prunella vulgaris</i> var. <i>elongata</i>	self-heal	Native	1
<i>Prunella vulgaris</i> var. <i>Vulgaris</i>	common selfheal	Introduced	0
<i>Prunus americana</i>	American plum	Native	3
<i>Prunus pensylvanica</i>	bird cherry	Native	3
<i>Prunus serotina</i>	black cherry	Native	0
<i>Pseudognaphalium obtusifolium</i>	sweet everlasting	Native	2
<i>Pycnanthemum tenuifolium</i>	slender mountain mint	Native	7
<i>Pycnanthemum virginianum</i>	common mountain mint	Native	5
<i>Ratibida pinnata</i>	gray-headed coneflower	Native	4
<i>Rhamnus cathartica</i>	common buckthorn	Introduced	0
<i>Rhus glabra</i>	smooth sumac	Native	1
<i>Rhus typhina</i>	staghorn sumac	Native	1
<i>Robinia pseudoacacia</i>	black locust	Introduced (N. America)	0
<i>Rosa carolina</i>	carolina rose	Native	5
<i>Rubus allegheniensis</i>	common blackberry	Native	3
<i>Rubus occidentalis</i>	black raspberry	Native	0
<i>Rudbeckia subtomentosa</i>	sweet black eyed susan	Native	8

<i>Rudbeckia triloba</i>	brown eyed susan	Native	1
<i>Ruellia humilis</i>	hairy wild petunia	Native	8
<i>Rumex crispus</i>	curled dock	Introduced	0
<i>Salix interior</i>	interior sandbar willow	Native	2
<i>Sanicula canadensis</i>	Canada black snakeroot	Native	5
<i>Schizachyrium scoparium</i>	little bluestem	Native	5
<i>Scirpus cyperinus</i>	woolgrass	Native	6
<i>Scutellaria parvula</i>	small skullcap	Native	8
<i>Securigera varia</i>	purple crownvetch	Introduced	0
<i>Setaria faberi</i>	giant foxtail	Introduced	0
<i>Silphium integrifolium</i>	prairie rosinweed	Native	5
<i>Silphium laciniatum</i>	compass plant	Native	5
<i>Silphium perfoliatum</i>	cup plant	Native	5
<i>Silphium terebinthinaceum</i>	prairie dock	Native	5
<i>Smilacina stellata</i>	starry Solomon's Seal	Native	5
<i>Smilax lasioneuron</i>	midwestern carrionflower	Native	5
<i>Solanum carolinense</i>	carolina horsenettle	Introduced (N. America)	0
<i>Solidago altissima</i>	tall goldenrod	Native	1
<i>Solidago canadensis</i>	Canada goldenrod	Native	1
<i>Solidago gigantea</i>	smooth goldenrod	Native	4
<i>Solidago juncea</i>	early goldenrod	Native	3
<i>Solidago nemoralis</i>	field goldenrod	Native	3
<i>Solidago patula</i>	swamp goldenrod	Native	9
<i>Solidago speciosa</i>	showy goldenrod	Native	8
<i>Sonchus arvensis</i>	perennial sowthistle	Introduced	0
<i>Sorghastrum nutans</i>	Indian grass	Native	5
<i>Spartina pectinata</i>	prairie cordgrass	Native	4
<i>Sporobolus heterolepsis</i>	prairie dropseed	Native	10
<i>Stachys hispida</i>	hairy hedgenettle	Native	5
<i>Stachys pilosa</i>	prairie woundwort	Native	5
<i>Steironema lanceolatum</i>	lanceleaved loosestrife	Native	9
<i>Symphyotrichum drummondii</i>	Drummond's aster	Native	3
<i>Symphyotrichum ericoides</i>	heath aster	Native	6
<i>Symphyotrichum firmum</i>	shining aster	Native	7
<i>Symphyotrichum laeve</i>	smooth blue aster	Native	9
<i>Symphyotrichum lanceolatum</i>	panicled aster	Native	3
<i>Symphyotrichum lateriflorum</i>	calico aster	Native	4
<i>Symphyotrichum novae-angliae</i>	New England aster	Native	3
<i>Symphyotrichum oblongifolium</i>	aromatic aster	Native	9
<i>Symphyotrichum oolentangiense</i>	sky blue aster	Native	8
<i>Symphyotrichum pilosum</i>	frost aster	Native	0
<i>Symphyotrichum praealtum</i>	willowleaf aster	Native	7
<i>Taraxacum officinale</i>	common dandelion	Introduced	0
<i>Tephrosia virginiana</i>	goat's rue	Native	8

<i>Thalictrum dasycarpum</i>	purple meadow-rue	Native	6
<i>Thalictrum dioicum</i>	early meadow rue	Native	7
<i>Trifolium hybridum</i>	alsike clover	Introduced	0
<i>Trifolium pratense</i>	red clover	Introduced	0
<i>Triosteum aurantiacum</i>	orange-fruited horse gentian	Native	5
<i>Ulmus americana</i>	American elm	Native	3
<i>Unidentified Asteraceae sp.</i>	Family; Out of season	Unknown	0
<i>Unidentified Asteraceae sp.</i>	unidentified basal leaves hairy entire margins	Unknown	0
<i>Unidentified Asteraceae sp.</i>	unknown aster family hairy entire basal leaves	Unknown	0
<i>Unidentified Asteraceae sp.</i>	unknown clasping aster family	Unknown	0
<i>Unidentified Cuscuta sp.</i>	unidentified dodder, growing on sunflower	Unknown	0
<i>Unidentified Geranium sp.</i>	unknown geranium	Unknown	0
<i>Unidentified grass</i>	thick, ornamental	Unknown	0
<i>Unidentified leaf 1</i>	<5%	Unknown	0
<i>Unidentified leaf 2</i>	<5%	Unknown	0
<i>Unidentified Mentha sp.</i>	unknown mint	Unknown	0
<i>Unidentified Poaceae sp.</i>	unknown true grass; out of season	Unknown	0
<i>Unidentified Poaceae sp.</i>	unknown thin true grass; possibly ornamental	Unknown	0
<i>Unidentified Poaceae sp.</i>	unknown true grass	Unknown	0
<i>Unidentified Solidago sp.</i>	unknown goldenrod genus rough serrate basal leaves	Unknown	0
<i>Unidentified sp.</i>	wavy textured basal leaves	Unknown	0
<i>Verbascum thapsus</i>	great mullein	Introduced	0
<i>Verbena hastata</i>	blue vervain	Native	4
<i>Verbena stricta</i>	hoary vervain	Native	4
<i>Verbena urticifolia</i>	white vervain	Native	2
<i>Vernonia missurica</i>	Missouri ironweed	Native	4
<i>Veronicastrum virginicum</i>	Culver's root	Native	8
<i>Viola pedatifida</i>	prairie violet	Native	9
<i>Viola sororia</i>	common blue violet	Native	3
<i>Vitis riparia</i>	riverbank grape	Native	1
<i>Zizia aptera</i>	heart-leaved golden Alexander	Native	10
<i>Zizia aurea</i>	golden alexander	Native	5

Appendix E: Information Regarding All Zones and Sites

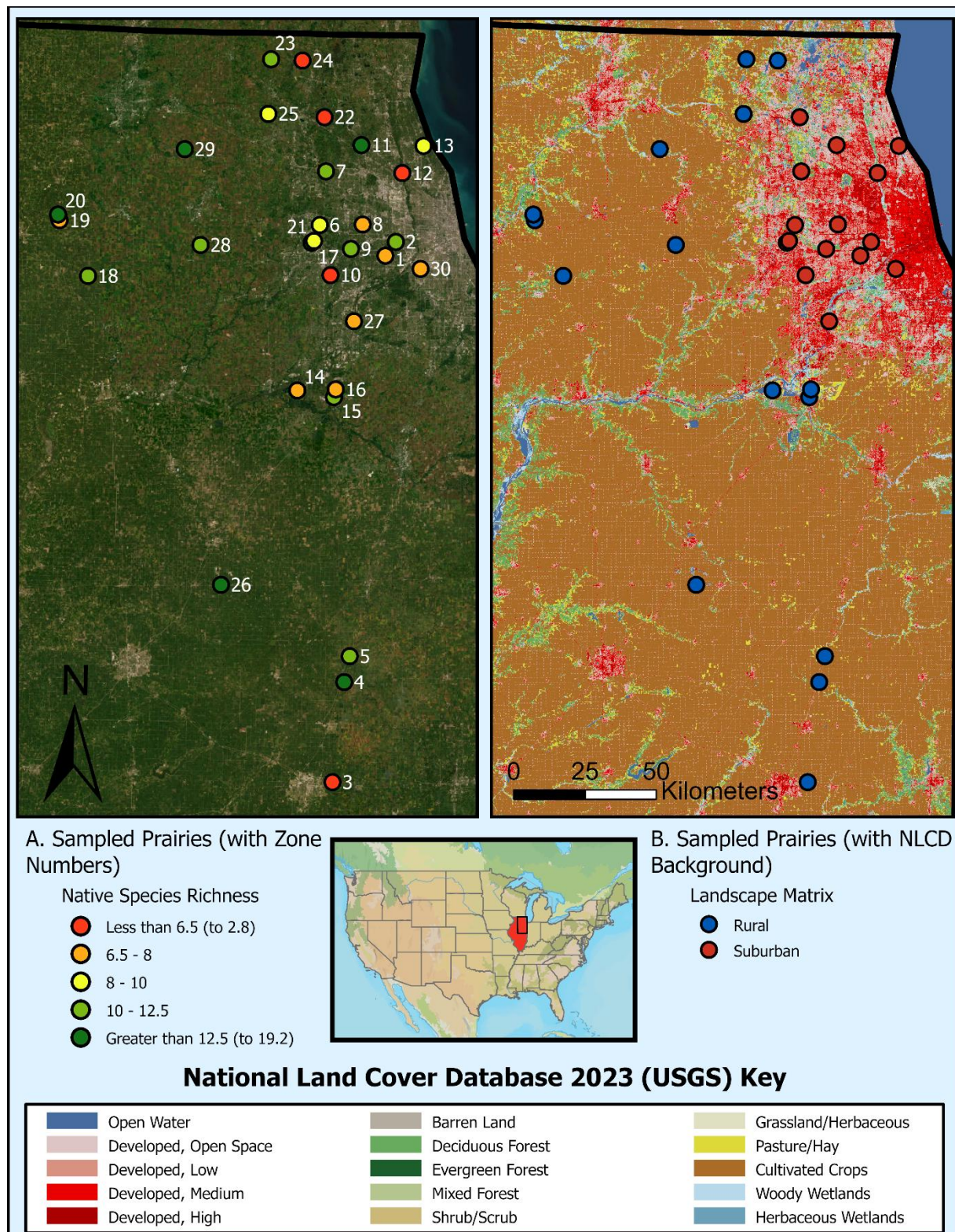


Figure E1 (Fig. 5 From Thesis). A map demonstrating the geographic distribution of all thirty sampled zones. The floral surveys collected from these thirty sites made up the backbone of the in-field primary dataset of this project. Zones were chosen from a list of potential prairies throughout northern and central Illinois, which was then refined to ensure that prairies were largely comparable in their ecology and management. Zones were selected evenly to guarantee equal rural and suburban samples.

Index Table of Zones, Sites, and Ownership

Zone	Zone Name	Landscape Matrix Class	Operating Organization
1	Hinsdale INAI Prairie	Suburban	County of DuPage
2	Wolf Road Prairie	Suburban	Forest Preserves of Cook County
3*	Trelease Prairie	Rural	University of Illinois
4	Prospect Cemetery**	Rural	Grand Prairie Friends
5	Loda Cemetery Prairie Nature Preserve**	Rural	Grand Prairie Friends
6	West Chicago Prairie	Suburban	Forest Preserve District of DuPage County
7	Shoe Factory Road Prairie	Suburban	Forest Preserves of Cook County
8	Churchill Woods Prairie	Suburban	Forest Preserve District of DuPage County
9	Morton Arboretum Schulenberg Prairie**	Suburban	Morton Arboretum
10	Springbrook Prairie	Suburban	Forest Preserve District of DuPage County
11	Deer Grove East	Suburban	Forest Preserves of Cook County
12	Kloempken Prairie	Suburban	Forest Preserves of Cook County
13	Chicago Botanic Garden Dixon Prairie	Suburban	Chicago Botanic Garden
14	Goose Lake Prairie	Rural	Illinois Department of Natural Resources
15	Midewin South Patrol Road	Rural	U.S. Forest Service
16	Midewin Glacial Plains	Rural	U.S. Forest Service
17	Fermilab Prairie Six	Suburban	U.S. Department of Energy
18	Richardson Wildlife Prairie**	Rural	Richardson Wildlife Foundation
19	Nachusa Thelma Carpenter	Rural	The Nature Conservancy
20	Nachusa Clear Creek 2006	Rural	The Nature Conservancy
21	Fermilab NE Circle	Suburban	U.S. Department of Energy
22	Hickory Grove Highlands	Suburban	McHenry County Conservation District
23	Hackmatack Queen Anne Prairie	Rural	U.S. Fish and Wildlife Service
24	Glacial Park South	Rural	McHenry County Conservation District
25	Pleasant Valley Conservation Area	Rural	McHenry County Conservation District
26	Weston Cemetery**	Rural	Parklands Foundation
27	Lockport Prairie	Suburban	Forest Preserve District of Will County
28	Afton South Prairie	Rural	DeKalb County Forest Preserve District
29	Nelson Prairie	Rural	DeKalb County Forest Preserve District
30	Bridgeview INAI Prairie	Suburban	Village of Bridgeview

*Zone 3 was dropped from most analyses as an outlier.

** Denotes private ownership.

Zone and Site designations are the same for all samples, except for the zones as follows: 15, 16 (Site 15: Midewin National Tallgrass Prairie); 17, 21 (Site 17: Fermilab National Laboratory); 19, 20 (Site 19: Nachusa Grasslands)

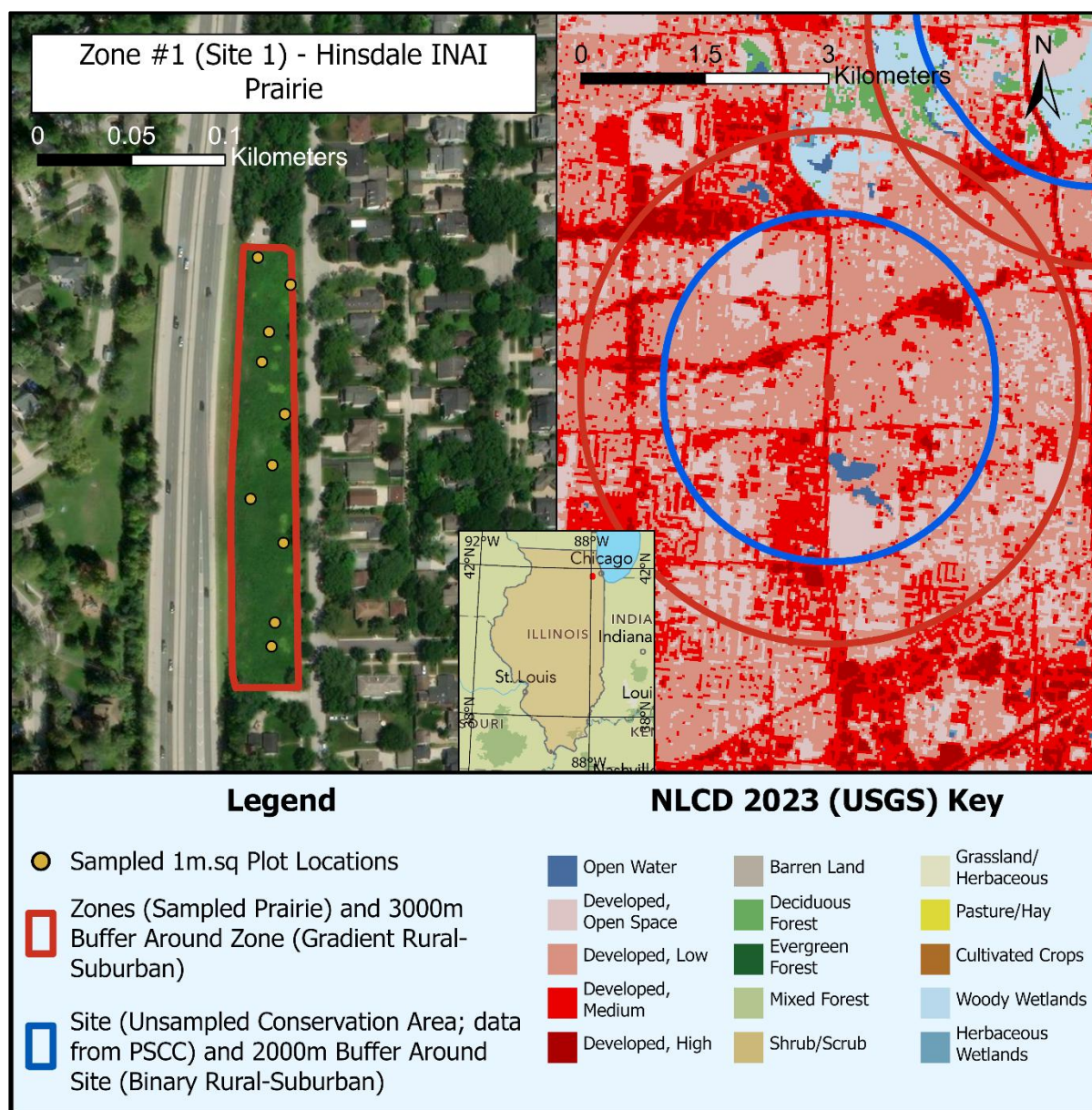
Index Table of Zone Scores, Ranked by Native FQI

Zone	Short Name	Native Sp. Richness	% Introduced	Native FQI	AFQI	Shannon Diversity (All)	Native Gamma Diversity
26	Weston	14.7	7.188	26.385	66.665	3.329	41
20	Nachusa Clear	16.6	6.398	24.198	57.473	3.325	40
29	Nelson	19.2	16.128	22.022	45.934	3.437	34
4	Prospect	14.3	0.714	19.576	51.359	3.294	39
11	Deer Grove	13.3	12.706	18.743	48.173	3.570	46
17	Fermilab Six	14.3	5.657	18.584	46.945	3.482	45
9	Morton	10.4	3.217	18.252	55.936	3.236	35
18	Richardson	12.1	4.577	17.002	47.701	3.141	32
7	Shoe Factory	11.2	4.807	16.441	48.579	3.313	39
15	Midewin South	12.2	6.251	16.219	45.584	3.301	37
28	Afton	10.1	16.471	15.907	45.895	3.339	30
5	Loda	10.6	5.965	15.197	45.041	3.275	34
13	CBG	9.6	4.644	15.189	48.612	3.522	44
23	Queen Anne	11.3	21.603	13.566	35.870	3.346	30
21	Fermilab NE	9.1	16.527	12.741	38.104	3.155	40
19	Nachusa Thelma	7.9	21.084	12.492	39.722	3.378	35
30	Bridgeview	7.9	4.270	11.723	41.683	2.927	27
27	Lockport	7.3	22.865	11.555	37.523	3.043	25
2	Wolf Road	10.1	1.769	11.433	34.509	3.222	38
14	Goose Lake	7.6	14.802	11.384	37.984	3.290	36
25	Pleasant Valley	8.3	16.862	10.851	34.046	3.190	33
8	Churchill	6.9	1.250	10.393	36.873	2.966	27
1	Hinsdale	6.7	3.790	10.340	37.476	3.159	30
6	West Chicago	8.6	11.726	10.014	31.650	3.179	34
16	Midewin Glacial	8.0	30.659	9.862	27.900	3.297	34
22	Hickory Grove	5.3	24.841	8.697	34.491	3.037	27
12	Kloempken	4.3	21.600	8.262	37.254	2.616	22
24	Glacial Park South	5.6	30.547	7.642	27.660	2.994	24
10	Springbrook	4.8	13.762	6.652	28.590	2.384	19
3*	Trelease	2.8	36.167	4.911	23.201	1.820	8

*Zone 3 was dropped from most analyses as an outlier.

Excepting Shannon diversity and gamma diversity, all values are the average of ten plot measurements.

Zone 1 (Site 1): Hinsdale INAI Prairie



Approximate Coordinates: 41°47'36"N 87°56'47"W

Date Sampled: August 20-21, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Full Remnant

Operating Organization: County of DuPage

Full Species List: *Andropogon gerardii*, *Bromus inermis*, *Ceanothus americanus*, *Coreopsis palmata*, *Coreopsis tripteris*, *Cornus racemosa*, *Cornus sericea*, *Eryngium yuccifolium*, *Eupatorium altissimum*, *Euthamia graminifolia*, *Fragaria virginiana*, *Geum canadense*, *Helianthus divaricatus*, *Helianthus roseserratus*, *Juglans nigra*, *Liatris aspera*, *Monarda fistulosa*, *Oenothera biennis*, *Oligoneuron rigidum*, *Panicum capillare*, *Parthenium integrifolium*, *Parthenocissus inserta*, *Phlox paniculata*, *Pycnanthemum tenuifolium*, *Rosa carolina*, *Rudbeckia triloba*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium terebinthinaceum*, *Smilacina stellata*, *Solidago altissima*, Unidentified grass (thick, ornamental), Unidentified leaf 1 (<5%), Unidentified leaf 2 (<5%)



Native Species Richness:	6.700
All Species Richness:	7.500
Introduced Species Percent:	3.790
Mean C (Native):	4.001
Mean C (All):	3.776
FQI (Native):	10.340
FQI (All):	9.760
AFQI:	37.476
Shannon Diversity (All):	3.159
Simpson Diversity (All):	0.946
Native Gamma Diversity:	30
All Gamma Diversity:	34

Zone Area (ha):	0.75738
Site Area (with Zone):	0.75738
Distance to Chicago (km):	22.1549
Avg. % Rural (8 Buffers):	0
Avg. % Suburban:	98.8446
Avg. % Wilderness:	0.21215
Avg. Ratio of Rural to Suburban (Log):	Undefined
Age of Restoration:	Remnant
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization	
Budget Class:	Undefined
Ownership Type (PSCC):	Municipal

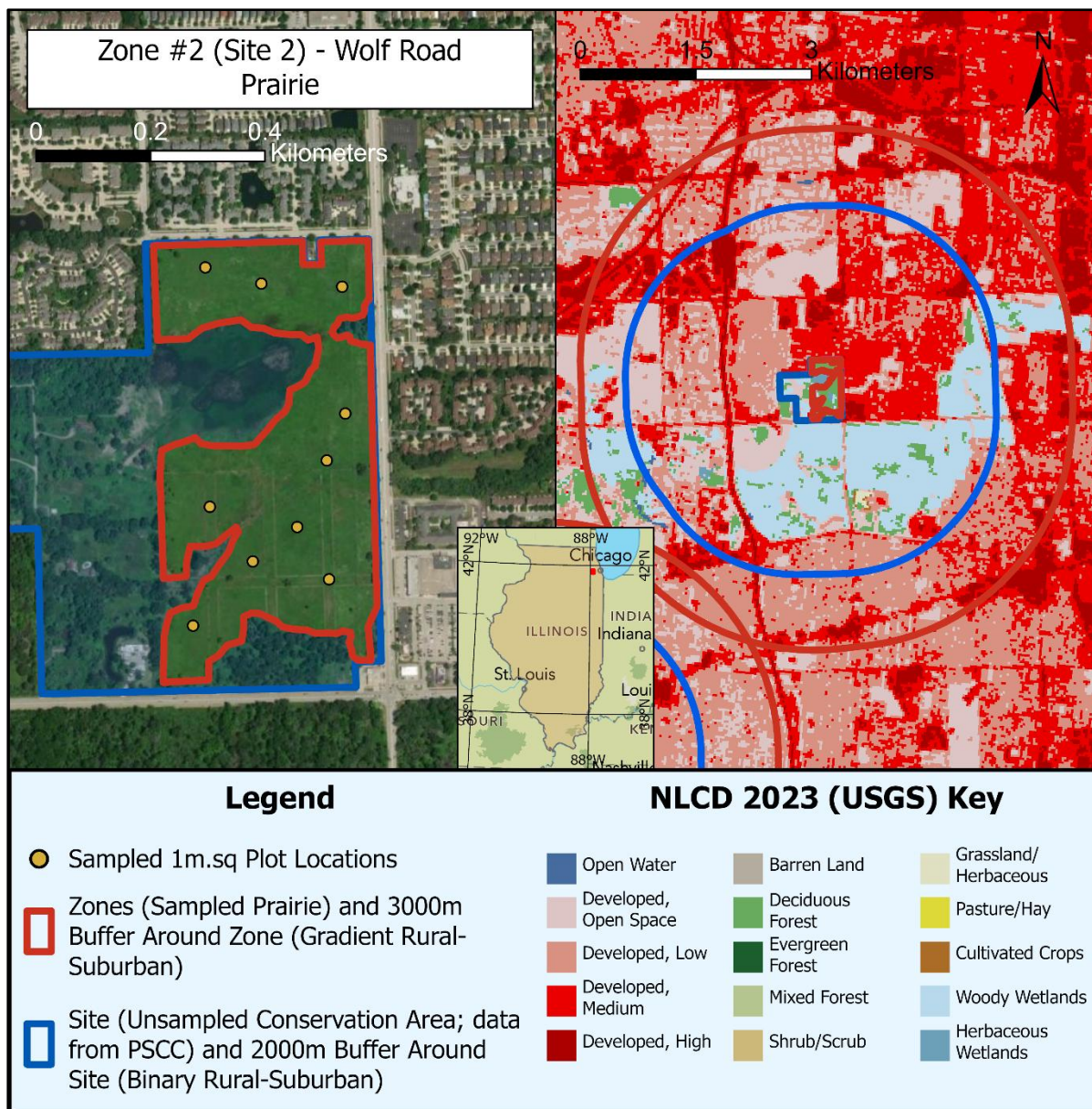
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Coreopsis palmata*, *Eryngium yuccifolium*, *Parthenium integrifolium*, *Liatris aspera*, *Ceanothus americanus*

Top Five Highest Frequency/Coverage Native Species: *Silphium integrifolium*, *Solidago altissima*, *Silphium laciniatum*, *Rosa carolina*, *Monarda fistulosa*

Top Five Highest Frequency/Coverage All Species: *Bromus inermis*, *Unidentified grass*, *Silphium integrifolium*, *Solidago altissima*, *Silphium laciniatum*

Zone 2 (Site 2): Wolf Road Prairie



Approximate Coordinates: 41°50'12"N 87°54'13"W

Date Sampled: August 22, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Mixed

Operating Organization: Forest Preserves of Cook County

Notes of Interest: Originally planned as a housing development in the 1920's. It never happened.

Full Species List: *Achillea millefolium*, *Agrimonia gryposepala*, *Apocynum cannabinum*, *Blephilia hirsuta*, *Coreopsis tripteris*, *Cornus racemosa*, *Erigeron annuus*, *Eryngium yuccifolium*, *Eupatorium altissimum*, *Euphorbia corollata*, *Euthamia graminifolia*, *Fragaria virginiana*, *Gaura biennis*, *Geum aleppicum*, *Geum canadense*, *Helianthus grosseserratus*, *Hypericum punctatum*, *Liatris pycnostachya*, *Liatris spicata*, *Lycopus americanus*, *Monarda fistulosa*, *Oenothera biennis*, *Oligoneuron rigidum*, *Oxypolis rigidior*, *Parthenium integrifolium*, *Pseudognaphalium obtusifolium*, *Pycnanthemum tenuifolium*, *Pycnanthemum virginianum*, *Rhamnus cathartica*, *Rosa carolina*, *Rubus allegheniensis*, *Rubus occidentalis*, *Schizachyrium scoparium*, *Silphium integrifolium*, *Silphium terebinthinaceum*, *Solidago altissima*, *Solidago canadensis*, *Symphotrichum ericoides*, *Unidentified Geranium sp.*, *Unidentified Mentha sp.*, *Unidentified Poaceae sp.*, *Vernonia missurica*, *Viola sororia*



Native Species Richness:	10.100
All Species Richness:	11.100
Introduced Species Percent:	1.769
Mean C (Native):	3.605
Mean C (All):	3.553
FQI (Native):	11.433
FQI (All):	11.268
AFQI:	34.509
Shannon Diversity (All):	3.222
Simpson Diversity (All):	0.937
Native Gamma Diversity:	38
All Gamma Diversity:	43

Zone Area (ha):	20.253
Site Area (with Zone):	48.4256
Distance to Chicago (km):	17.9904
Avg. % Rural (8 Buffers):	0.03757
Avg. % Suburban:	70.0505
Avg. % Wilderness:	29.862
Avg. Ratio of Rural to Suburban (Log):	-7.5935
Age of Restoration:	1982
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	> \$150m
Ownership Type (PSCC):	County

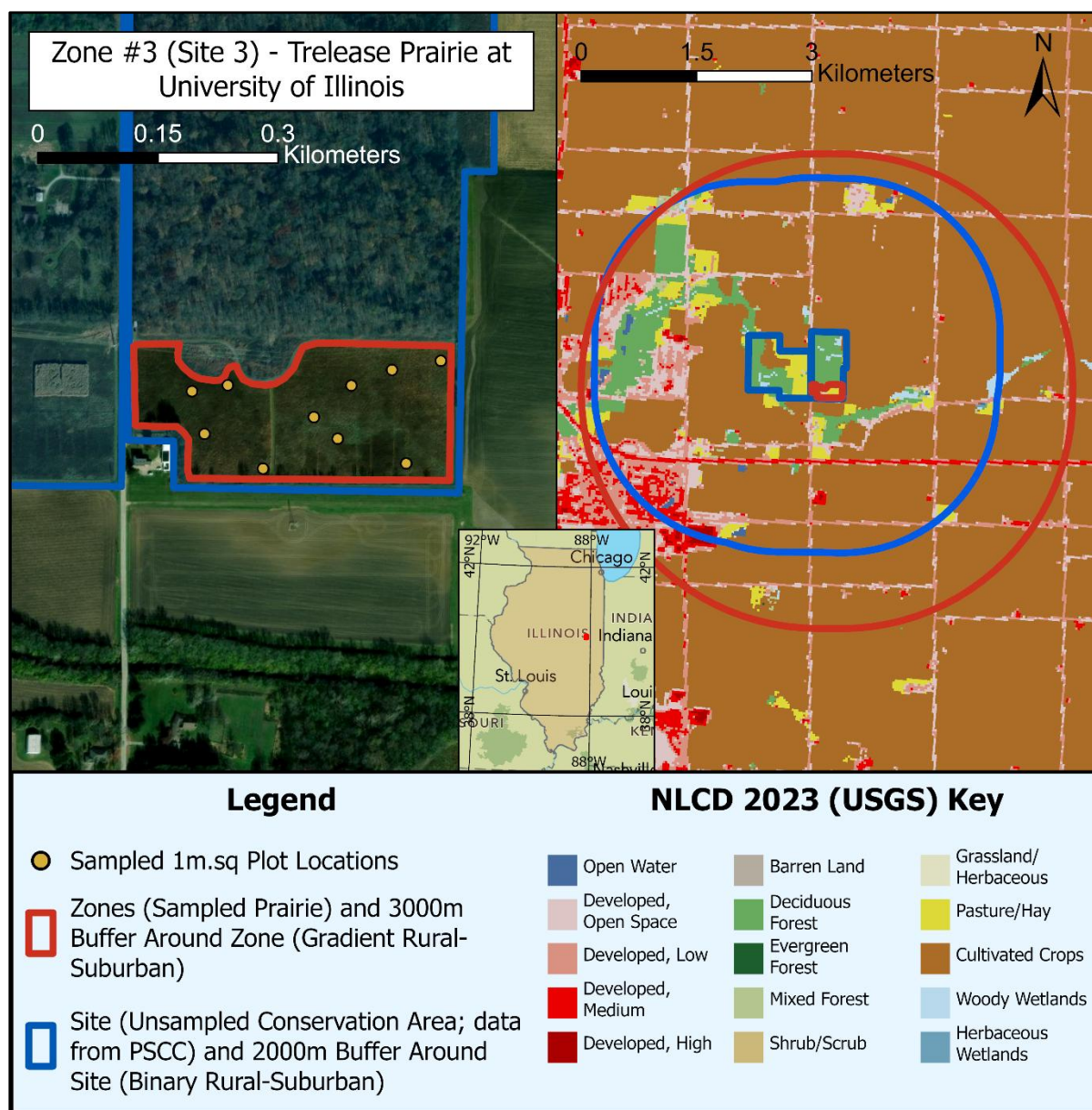
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Eryngium yuccifolium*, *Parthenium integrifolium*, *Oxypolis rigidior*, *Liatris pycnostachya*, *Pycnanthemum tenuifolium*

Top Five Highest Frequency/Coverage Native Species: *Solidago altissima*, *Euthamia graminifolia*, *Cornus racemosa*, *Monarda fistulosa*, *Fragaria virginiana*

Top Five Highest Frequency/Coverage All Species: *Solidago altissima*, *Euthamia graminifolia*, *Cornus racemosa*, *Monarda fistulosa*, *Fragaria virginiana*

Zone 3 (Site 3): Trelease Prairie



Approximate Coordinates: 40°07'42"N 88°08'25"W

Date Sampled: August 23, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Restoration

Operating Organization: University of Illinois at Urbana-Champaign

Full Species List: *Agrostis gigantea*, *Andropogon gerardii*, *Asclepias syriaca*, *Bromus inermis*, *Eupatorium altissimum*, *Lespedeza cuneata*, *Monarda fistulosa*, *Pycnanthemum tenuifolium*, *Silphium integrifolium*, *Solidago altissima*, *Sorghastrum nutans*



Native Species Richness:	2.800
All Species Richness:	4.300
Introduced Species Percent:	36.167
Mean C (Native):	2.817
Mean C (All):	1.928
FQI (Native):	4.911
FQI (All):	3.429
AFQI:	23.201
Shannon Diversity (All):	1.820
Simpson Diversity (All):	0.772
Native Gamma Diversity:	8
All Gamma Diversity:	11

Zone Area (ha):	5.59892
Site Area (with Zone):	91.345
Distance to Chicago (km):	193.54
Avg. % Rural (8 Buffers):	59.1041
Avg. % Suburban:	13.2751
Avg. % Wilderness:	16.942
Avg. Ratio of Rural to Suburban (Log):	1.66105
Age of Restoration:	1945
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization	
Budget Class:	> \$150m
Ownership Type (PSCC):	State

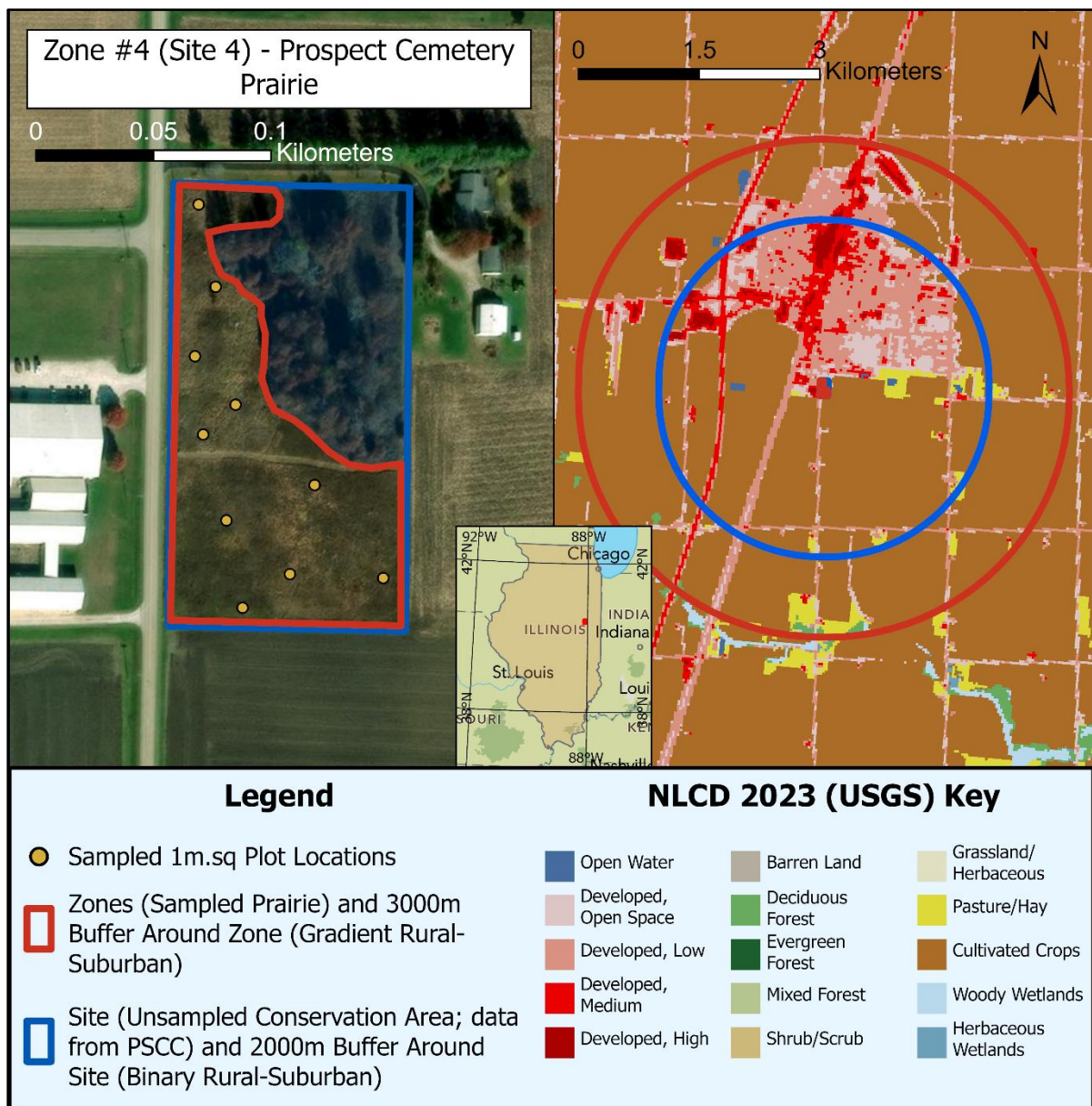
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Pycnanthemum tenuifolium*, *Sorghastrum nutans*, *Silphium integrifolium*, *Andropogon gerardii*, *Monarda fistulosa*

Top Five Highest Frequency/Coverage Native Species: *Solidago altissima*, *Andropogon gerardii*, *Monarda fistulosa*, *Sorghastrum nutans*, *Pycnanthemum tenuifolium*

Top Five Highest Frequency/Coverage All Species: *Solidago altissima*, *Lespedeza cuneata*, *Andropogon gerardii*, *Bromus inermis*, *Monarda fistulosa*

Zone 4 (Site 4): Prospect Cemetery Prairie



Approximate Coordinates: 40°26'42"N 88°05'49"W

Date Sampled: August 23-24, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Remnant

Operating Organization: Grand Prairie Friends

Full Species List: *Amorpha canescens*, *Andropogon gerardii*, *Anemone cylindrica*, *Arnoglossum lantagineum*, *Calystegia sepium*, *Ceanothus americanus*, *Celastrus scandens*, *Comandra umbellata*, *Coreopsis tripteris*, *Dichanthelium oligosanthes*, *Eryngium yuccifolium*, *Eupatorium altissimum*, *Euphorbia corollata*, *Fragaria virginiana*, *Gaura biennis*, *Helianthus grosseserratus*, *Helianthus mollis*, *Helianthus pauciflorus*, *Monarda fistulosa*, *Nabalus asper*, *Oligoneuron rigidum*, *Parthenium integrifolium*, *Pedicularis canadensis*, *Physostegia virginiana*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rosa carolina*, *Ruellia humilis*, *Sanicula canadensis*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium terebinthinaceum*, *Solidago altissima*, *Sorghastrum nutans*, *Steironema lanceolatum*, *Symphyotrichum ericoides*, *Symphyotrichum novae-angliae*, *Unidentified Asteraceae sp.*, *Unidentified Poaceae sp.*, *Viola pedatifida*, *Viola sororia*, *Vitis riparia*



Native Species Richness:	14.300
All Species Richness:	14.600
Introduced Species Percent:	0.714
Mean C (Native):	5.202
Mean C (All):	5.159
FQI (Native):	19.576
FQI (All):	19.421
AFQI:	51.359
Shannon Diversity (All):	3.294
Simpson Diversity (All):	0.948
Native Gamma Diversity:	39
All Gamma Diversity:	42

Zone Area (ha):	1.09772
Site Area (with Zone):	1.8992
Distance to Chicago (km):	158.395
Avg. % Rural (8 Buffers):	55.9649
Avg. % Suburban:	42.285
Avg. % Wilderness:	0.03978
Avg. Ratio of Rural to Suburban (Log):	0.36155
Age of Restoration:	Remnant
Sector Type (PSCC):	Private
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	< \$5m
Ownership Type (PSCC):	NGO

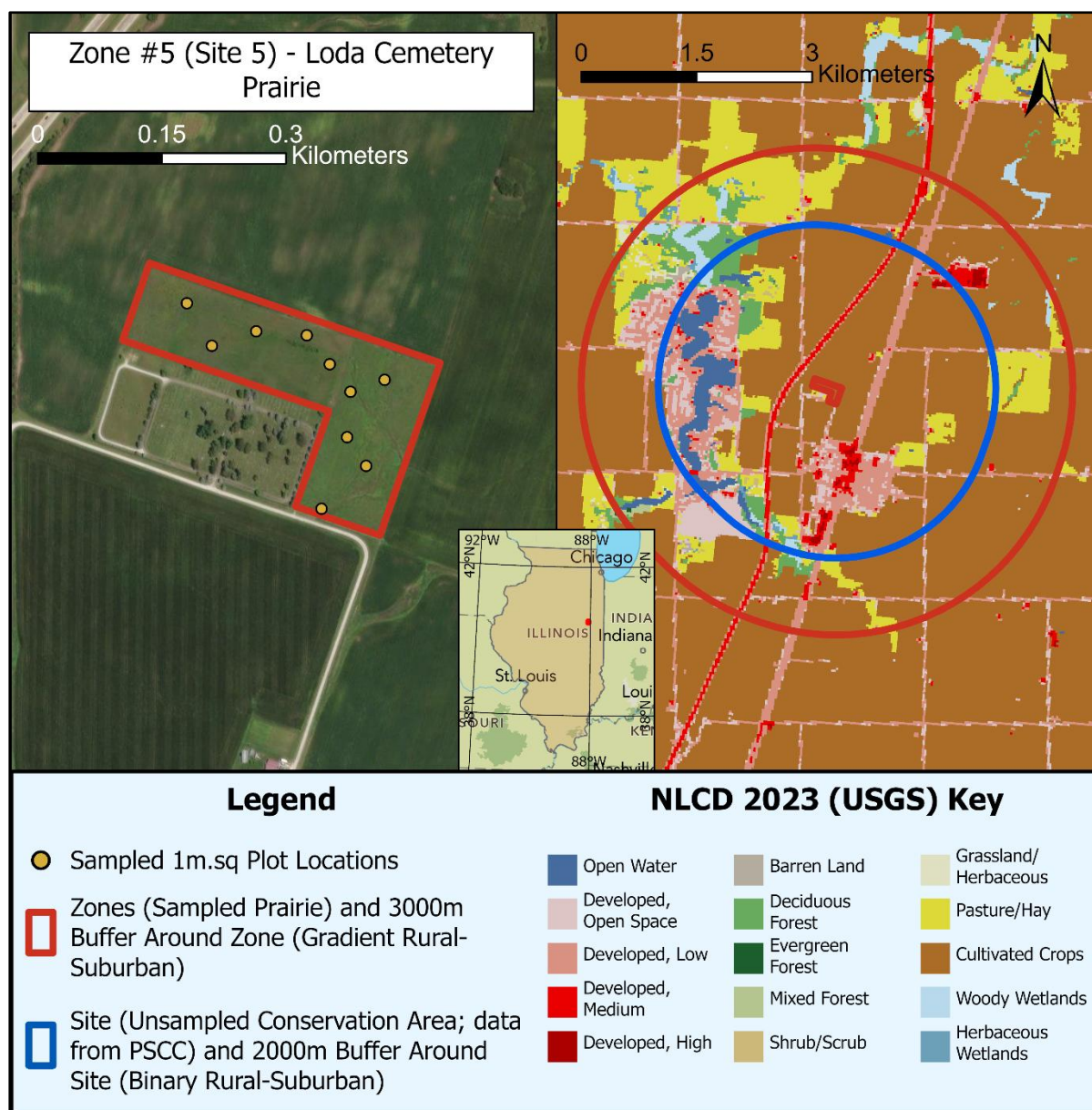
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Nabalus asper*, *Arnoglossum plantagineum*, *Amorpha canescens*, *Viola pedatifida*, *Steironema lanceolatum*,

Top Five Highest Frequency/Coverage Native Species: *Sorghastrum nutans*, *Solidago altissima*, *Symphyotrichum ericoides*, *Ceanothus americanus*, *Andropogon gerardii*,

Top Five Highest Frequency/Coverage All Species: *Sorghastrum nutans*, *Solidago altissima*, *Symphyotrichum ericoides*, *Ceanothus americanus*, *Andropogon gerardii*,

Zone 5 (Site 5): Loda Cemetery Prairie



Approximate Coordinates: 40°31'35"N 88°04'27"W

Date Sampled: August 24, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Mixed

Operating Organization: Grand Prairie Friends

Full Species List: *Ambrosia artemisiifolia*, *Amorpha canescens*, *Andropogon gerardii*, *Apocynum cannabinum*, *Baptisia australis*, *Chamaecrista fasciculata*, *Coreopsis tripteris*, *Dalea purpurea*, *Elymus canadensis*, *Eryngium yuccifolium*, *Euthamia graminifolia*, *Fragaria virginiana*, *Helianthus rosseserratus*, *Helianthus pauciflorus*, *Lycopus americanus*, *Melilotus officinalis*, *Monarda fistulosa*, *Panicum capillare*, *Parthenium integrifolium*, *Potentilla simplex*, *Prunus americana*, *Pycnanthemum tenuifolium*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rubus allegheniensis*, *Rudbeckia subtomentosa*, *Ruellia humilis*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium terebinthinaceum*, *Solidago altissima*, *Sorghastrum nutans*, *Symphotrichum ericoides*, *Symphotrichum novae-angliae*, *Symphotrichum pilosum*, *Trifolium pratense*, *Zizia aptera*



Native Species Richness:	10.600
All Species Richness:	11.200
Introduced Species Percent:	5.965
Mean C (Native):	4.637
Mean C (All):	4.380
FQI (Native):	15.197
FQI (All):	14.437
AFQI:	45.041
Shannon Diversity (All):	3.275
Simpson Diversity (All):	0.953
Native Gamma Diversity:	34
All Gamma Diversity:	37

Zone Area (ha):	4.98649
Site Area (with Zone):	4.98649
Distance to Chicago (km):	149.112
Avg. % Rural (8 Buffers):	66.3374
Avg. % Suburban:	21.2038
Avg. % Wilderness:	2.51362
Avg. Ratio of Rural to Suburban (Log):	1.29238
Age of Restoration:	2008
Sector Type (PSCC):	Private
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	< \$5m
Ownership Type (PSCC):	NGO

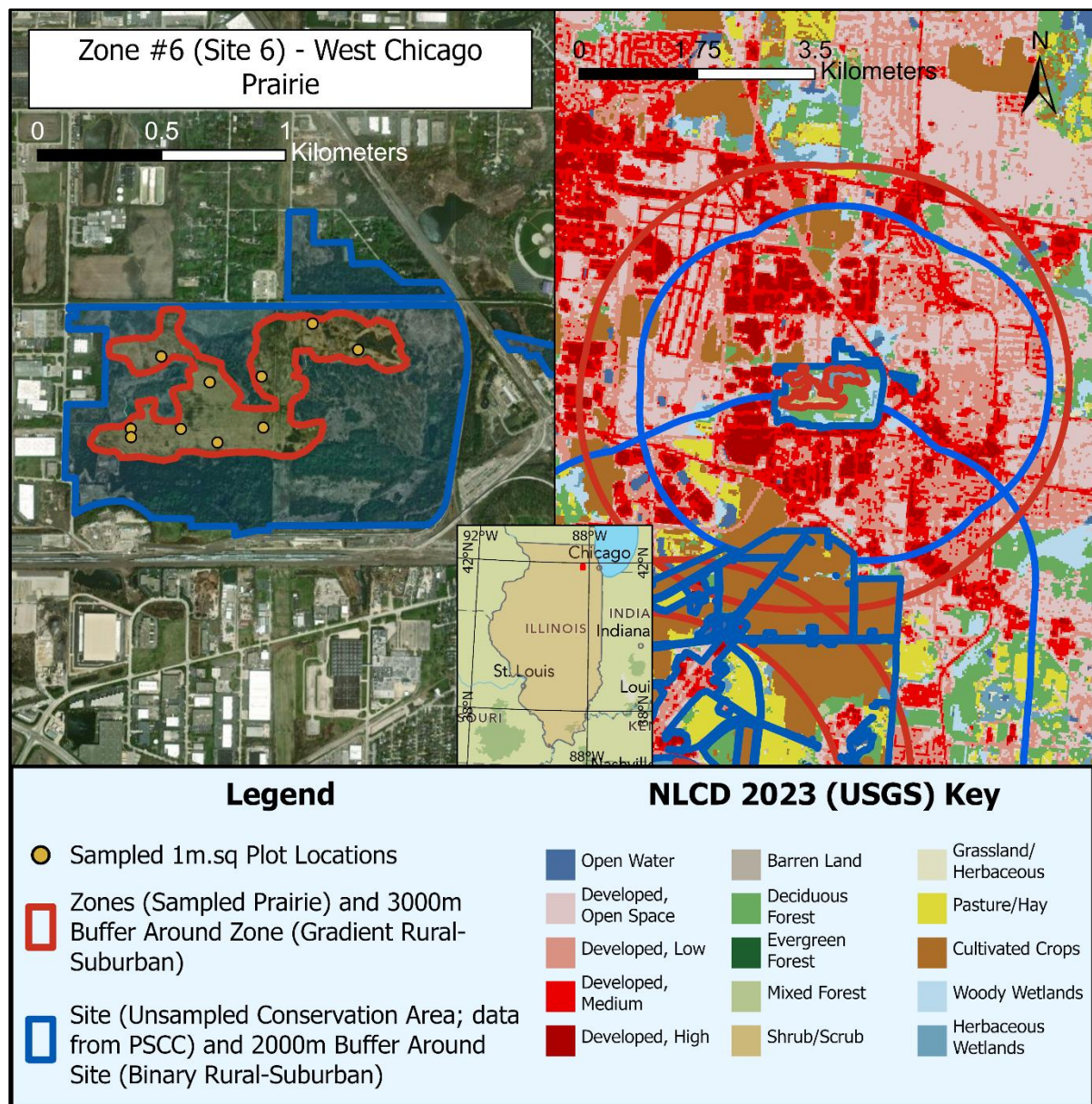
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Zizia aptera*, *Amorpha canescens*, *Helianthus pauciflorus*, *Eryngium yuccifolium*, *Dalea purpurea*,

Top Five Highest Frequency/Coverage Native Species: *Sorghastrum nutans*, *Elymus canadensis*, *Fragaria virginiana*, *Chamaecrista fasciculata*, *Solidago altissima*,

Top Five Highest Frequency/Coverage All Species: *Sorghastrum nutans*, *Elymus canadensis*, *Fragaria virginiana*, *Chamaecrista fasciculata*, *Solidago altissima*,

Zone 6 (Site 6): West Chicago Prairie



Approximate Coordinates: 41°53'18"N 88°13'37"W

Date Sampled: August 26, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Mixed

Operating Organization: Forest Preserve District of DuPage County

Full Species List: *Achillea millefolium*, *Agrimonia parviflora*, *Agrostis gigantea*, *Andropogon gerardii*, *Apios americana*, *Astragalus canadensis*, *Bromus inermis*, *Cirsium discolor*, *Coreopsis tripteris*, *Cornus racemosa*, *Daucus carota*, *Epilobium coloratum*, *Euthamia graminifolia*, *Fragaria virginiana*, *Gentiana andrewsii*, *Helianthus grosseserratus*, *Melilotus officinalis*, *Monarda fistulosa*, *Panicum capillare*, *Penstemon digitalis*, *Physostegia virginiana*, *Prunella vulgaris* var. *elongata*, *Prunus pensylvanica*, *Pycnanthemum tenuifolium*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rhamnus cathartica*, *Rhus typhina*, *Rosa carolina*, *Rubus allegheniensis*, *Silphium integrifolium*, *Silphium terebinthinaceum*, *Smilacina stellata*, *Solidago altissima*, *Sorghastrum nutans*, *Symphytotrichum pilosum*, *Trifolium hybridum*, *Unidentified Geranium* sp., *Veronicastrum virginicum*, *Vitis riparia*, *Zizia aurea*



Native Species Richness:	8.600
All Species Richness:	10.000
Introduced Species Percent:	11.726
Mean C (Native):	3.380
Mean C (All):	2.998
FQI (Native):	10.014
FQI (All):	8.917
AFQI:	31.650
Shannon Diversity (All):	3.179
Simpson Diversity (All):	0.939
Native Gamma Diversity:	34
All Gamma Diversity:	41

Zone Area (ha):	35.5062
Site Area (with Zone):	144.806
Distance to Chicago (km):	45.1017
Avg. % Rural (8 Buffers):	5.60114
Avg. % Suburban:	60.609
Avg. % Wilderness:	32.5398
Avg. Ratio of Rural to Suburban (Log):	-2.3856
Age of Restoration:	1995
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	\$5m-
Budget Class:	\$150m
Ownership Type (PSCC):	County

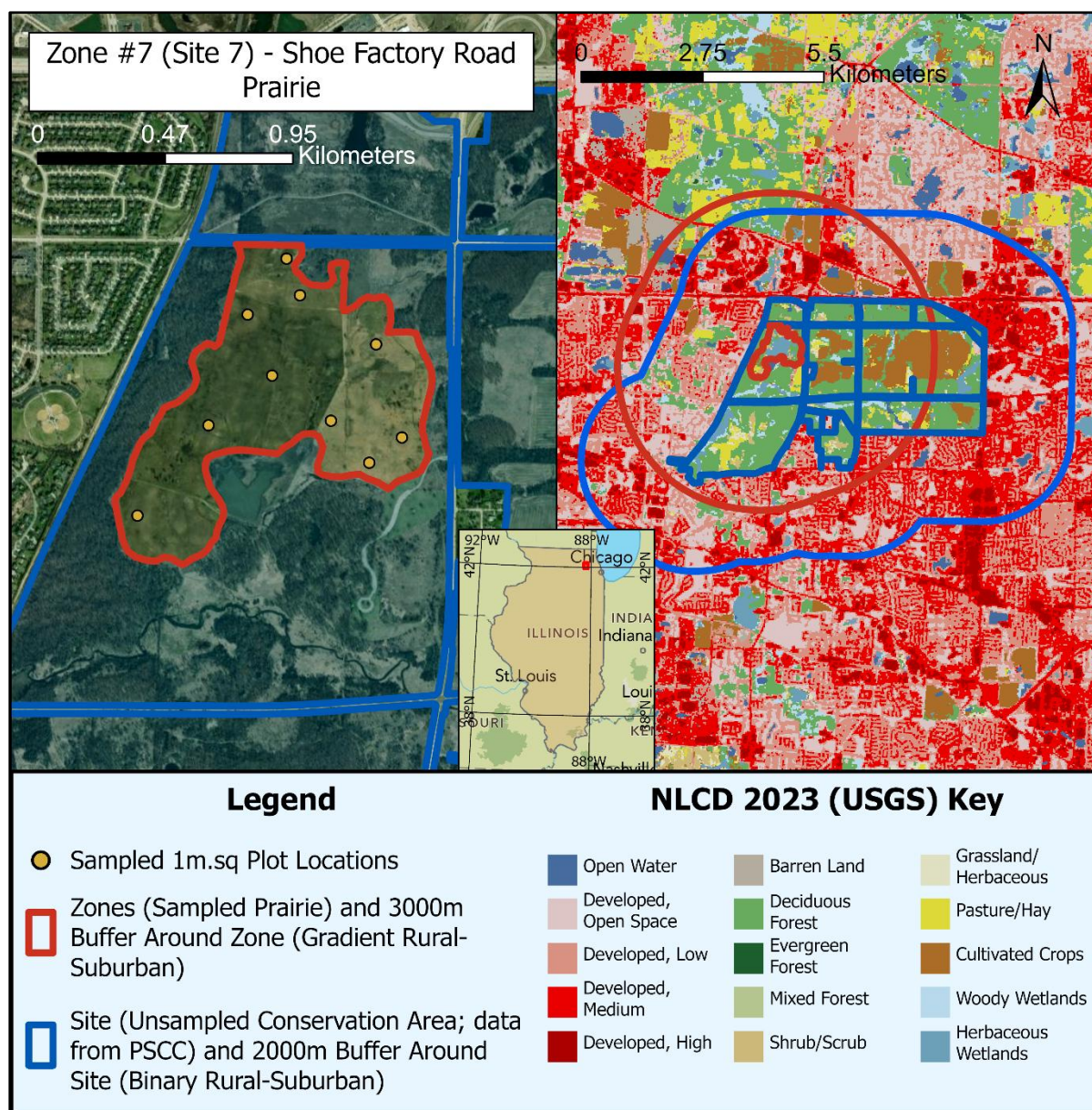
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Gentiana andrewsii*, *Veronicastrum virginicum*, *Astragalus canadensis*, *Pycnanthemum tenuifolium*, *Zizia aurea*,

Top Five Highest Frequency/Coverage Native Species: *Andropogon gerardii*, *Cornus racemosa*, *Sorghastrum nutans*, *Solidago altissima*, *Rosa carolina*,

Top Five Highest Frequency/Coverage All Species: *Andropogon gerardii*, *Cornus racemosa*, *Sorghastrum nutans*, *Solidago altissima*, *Agrostis gigantea*,

Zone 7 (Site 7): Shoe Factory Road Prairie



Approximate Coordinates: 42°03'23"N 88°11'53"W

Date Sampled: August 27, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Mixed

Operating Organization: Forest Preserves of Cook County

Full Species List: *Achillea millefolium*, *Agrostis gigantea*, *Amorpha canescens*, *Andropogon gerardii*, *Apocynum cannabinum*, *Coreopsis palmata*, *Coreopsis tripteris*, *Cornus racemosa*, *Dichanthelium oligosanthes*, *Echinacea pallida*, *Eryngium yuccifolium*, *Fragaria virginiana*, *Helianthus pauciflorus*, *Heliopsis helianthoides*, *Lespedeza capitata*, *Lespedeza cuneata*, *Liatris aspera*, *Monarda fistulosa*, *Oligoneuron ohioense*, *Oligoneuron rigidum*, *Oxypolis rigidior*, *Panicum capillare*, *Parthenium integrifolium*, *Pedicularis canadensis*, *Penthorum sedoides*, *Physostegia virginiana*, *Potentilla recta*, *Pycnanthemum tenuifolium*, *Ratibida pinnata*, *Rhamnus cathartica*, *Robinia pseudoacacia*, *Rubus occidentalis*, *Schizachyrium scoparium*, *Silphium integrifolium*, *Silphium laciniatum*, *Solidago altissima*, *Solidago canadensis*, *Solidago nemoralis*, *Solidago patula*, *Solidago speciosa*, *Sorghastrum nutans*, *Symphytotrichum ericoides*, *Symphytotrichum pilosum*, *Vitis riparia*



Native Species Richness:	11.200
All Species Richness:	11.800
Introduced Species Percent:	4.807
Mean C (Native):	4.982
Mean C (All):	4.743
FQI (Native):	16.441
FQI (All):	15.667
AFQI:	48.579
Shannon Diversity (All):	3.313
Simpson Diversity (All):	0.947
Native Gamma Diversity:	39
All Gamma Diversity:	44

Zone Area (ha):	70.7422
Site Area (with Zone):	1755.53
Distance to Chicago (km):	49.0642
Avg. % Rural (8 Buffers):	7.6367
Avg. % Suburban:	35.2509
Avg. % Wilderness:	51.9131
Avg. Ratio of Rural to Suburban (Log):	-1.311
Age of Restoration:	1990
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	> \$150m
Ownership Type (PSCC):	County

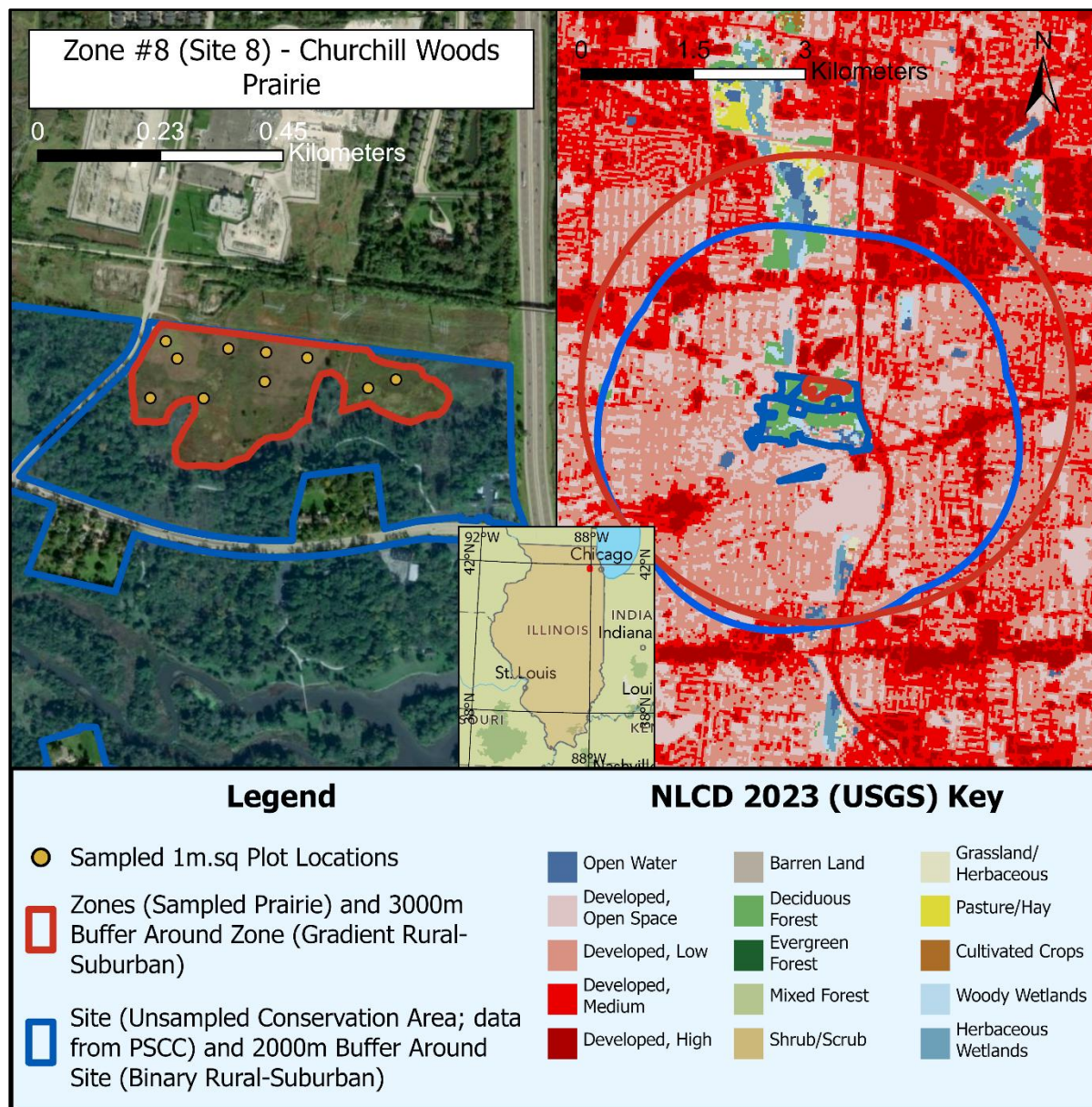
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Oligoneuron ohioense*, *Echinacea pallida*, *Coreopsis palmata*, *Amorpha canescens*, *Solidago patula*,

Top Five Highest Frequency/Coverage Native Species: *Sorghastrum nutans*, *Andropogon gerardii*, *Monarda fistulosa*, *Ratibida pinnata*, *Solidago altissima*,

Top Five Highest Frequency/Coverage All Species: *Sorghastrum nutans*, *Andropogon gerardii*, *Monarda fistulosa*, *Ratibida pinnata*, *Solidago altissima*,

Zone 8 (Site 8): Churchill Woods Prairie



Approximate Coordinates: 41°53'29"N 88°02'34"W

Date Sampled: August 28, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Mixed

Operating Organization: Forest Preserve District of DuPage County

Full Species List: *Achillea millefolium*, *Andropogon gerardii*, *Astragalus canadensis*, *Coreopsis tripteris*, *Cornus racemosa*, *Echinacea pallida*, *Eryngium yuccifolium*, *Eupatorium altissimum*, *Euthamia graminifolia*, *Fragaria virginiana*, *Helianthus grosseserratus*, *Lespedeza cuneata*, *Lycopus uniflorus*, *Monarda fistulosa*, *Parthenocissus inserta*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rosa carolina*, *Rubus allegheniensis*, *Silphium integrifolium*, *Silphium laciniatum*, *Solidago altissima*, *Solidago gigantea*, *Solidago juncea*, *Sorghastrum nutans*, *Symphyotrichum ericoides*, *Triosteum aurantiacum*, *Unidentified Asteraceae sp.*, *Unidentified Geranium sp.*, *Unidentified Poaceae sp.*, *Unidentified Solidago sp.*, *Veronicastrum virginicum*



Native Species Richness:	6.900
All Species Richness:	7.900
Introduced Species Percent:	1.250
Mean C (Native):	3.959
Mean C (All):	3.888
FQI (Native):	10.393
FQI (All):	10.204
AFQI:	36.873
Shannon Diversity (All):	2.966
Simpson Diversity (All):	0.928
Native Gamma Diversity:	27
All Gamma Diversity:	32

Zone Area (ha):	8.26969
Site Area (with Zone):	104.179
Distance to Chicago (km):	30.2408
Avg. % Rural (8 Buffers):	0.0819
Avg. % Suburban:	73.1513
Avg. % Wilderness:	25.9766
Avg. Ratio of Rural to Suburban (Log):	-6.5056
Age of Restoration:	1974
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization Budget Class:	\$5m-\$150m
Ownership Type (PSCC):	County

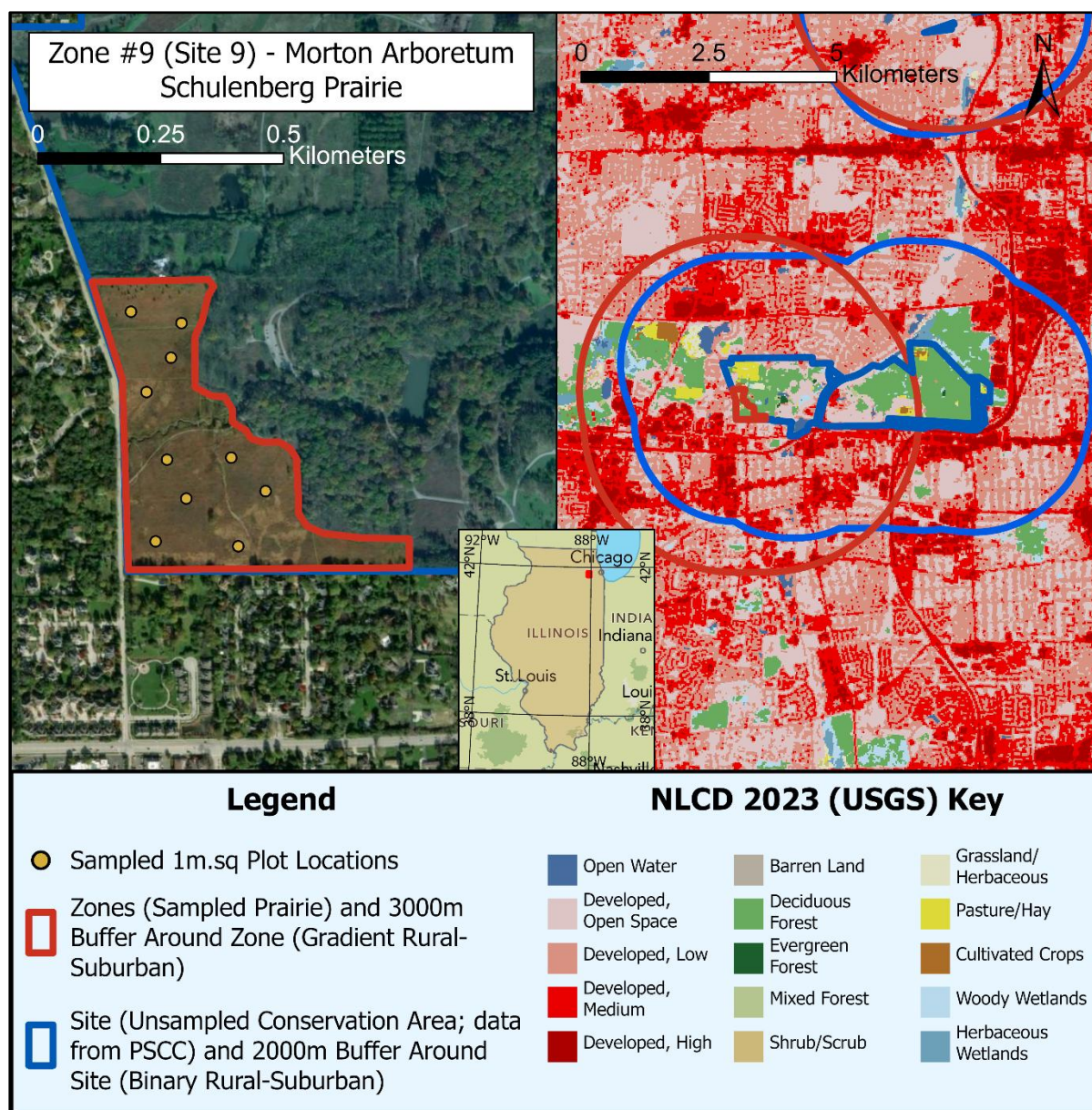
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Echinacea pallida*, *Eryngium yuccifolium*, *Veronicastrum virginicum*, *Astragalus canadensis*, *Symphotrichum ericoides*,

Top Five Highest Frequency/Coverage Native Species: *Andropogon gerardii*, *Sorghastrum nutans*, *Solidago altissima*, *Monarda fistulosa*, *Cornus racemosa*,

Top Five Highest Frequency/Coverage All Species: *Andropogon gerardii*, *Sorghastrum nutans*, *Solidago altissima*, *Monarda fistulosa*, *Solidago sp. unknown goldenron*

Zone 9 (Site 9): Morton Arboretum Schulenberg Prairie



Approximate Coordinates: 41°48'55"N 88°05'34"W

Date Sampled: August 29, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Fully Restoration

Operating Organization: Morton Arboretum

Full Species List: *Achillea millefolium*, *Agrostis gigantea*, *Amorpha canescens*, *Andropogon gerardii*, *Baptisia alba*, *Baptisia bracteata*, *Bromus inermis*, *Coreopsis palmata*, *Coreopsis tripteris*, *Cornus racemosa*, *Dalea candida*, *Dalea purpurea*, *Desmodium canadense*, *Dichanthelium oligosanthos*, *Eryngium yuccifolium*, *Fragaria virginiana*, *Galium mollugo*, *Gentiana alba*, *Helianthus mollis*, *Heuchera richardsonii*, *Lespedeza capitata*, *Monarda fistulosa*, *Oligoneuron rigidum*, *Panicum capillare*, *Parthenium integrifolium*, *Physostegia virginiana*, *Ratibida pinnata*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium terebinthinaceum*, *Solidago altissima*, *Solidago nemoralis*, *Sorghastrum nutans*, *Sporobolus heterolepsis*, *Symphyotrichum drummondii*, *Symphyotrichum ericoides*, *Symphyotrichum novae-angliae*, *Veronicastrum virginicum*



Native Species Richness:	10.400
All Species Richness:	10.800
Introduced Species Percent:	3.217
Mean C (Native):	5.656
Mean C (All):	5.538
FQI (Native):	18.252
FQI (All):	17.880
AFQI:	55.936
Shannon Diversity (All):	3.236
Simpson Diversity (All):	0.946
Native Gamma Diversity:	35
All Gamma Diversity:	38

Zone Area (ha):	17.0214
Site Area (with Zone):	591.324
Distance to Chicago (km):	33.7723
Avg. % Rural (8 Buffers):	0.67098
Avg. % Suburban:	72.8161
Avg. % Wilderness:	22.5396
Avg. Ratio of Rural to Suburban (Log):	-4.6439
Age of Restoration:	1962
Sector Type (PSCC):	Private
Special Legal Designations:	No
Operating Organization Budget Class:	\$5m-\$150m
Ownership Type (PSCC):	NGO

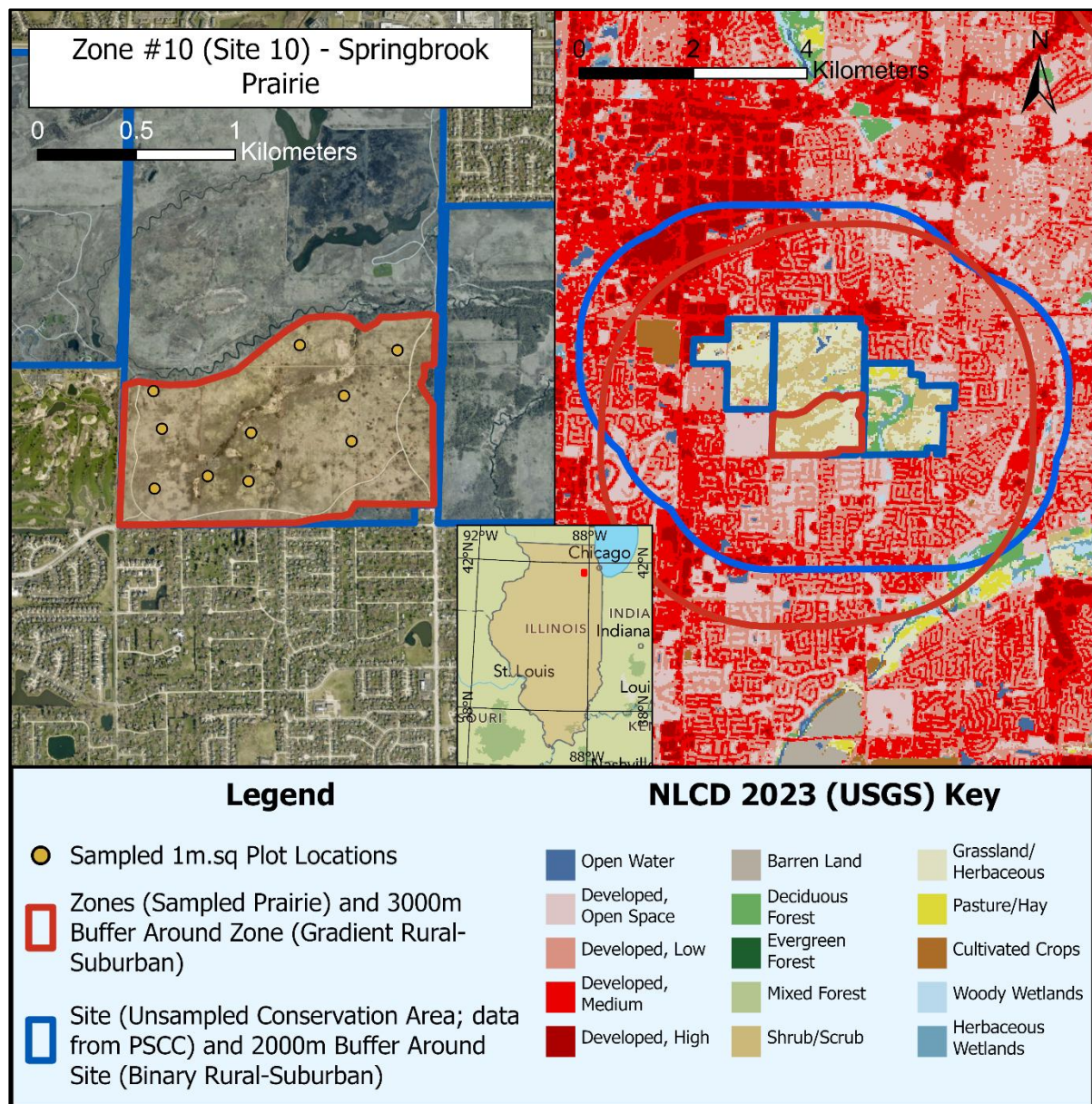
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Heuchera richardsonii* alumroot, *Dalea candida*, *Coreopsis palmata*, *Baptisia bracteata*,

Top Five Highest Frequency/Coverage Native Species: *Oligoneuron rigidum*, *Sorghastrum nutans*, *Andropogon gerardii*, *Dalea purpurea*, *Dalea candida*,

Top Five Highest Frequency/Coverage All Species: *Oligoneuron rigidum*, *Sorghastrum nutans*, *Andropogon gerardii*, *Dalea purpurea*, *Dalea candida*,

Zone 10 (Site 10): Springbrook Prairie



Approximate Coordinates: 41°44'16"N 88°10'40"W

Date Sampled: August 30, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Fully Restoration

Operating Organization: Forest Preserve District of DuPage County

Full Species List: *Andropogon gerardii*, *Asclepias verticillata*, *Bromus inermis*, *Chaerophyllum procumbens*, *Coreopsis tripteris*, *Cornus racemosa*, *Geum laciniatum*, *Hypericum punctatum*, *Monarda fistulosa*, *Oenothera biennis*, *Packera paupercula*, *Parthenocissus inserta*, *Pycnanthemum virginianum*, *Rosa carolina*, *Rubus occidentalis*, *Scirpus cyperinus*, *Silphium integrifolium*, *Solidago altissima*, *Sorghastrum nutans*, *Viola sororia*



Native Species Richness:	4.800
All Species Richness:	5.500
Introduced Species Percent:	13.762
Mean C (Native):	3.063
Mean C (All):	2.678
FQI (Native):	6.652
FQI (All):	5.892
AFQI:	28.590
Shannon Diversity (All):	2.384
Simpson Diversity (All):	0.868
Native Gamma Diversity:	19
All Gamma Diversity:	20

Zone Area (ha):	135.259
Site Area (with Zone):	737.653
Distance to Chicago (km):	42.4469
Avg. % Rural (8 Buffers):	0.41797
Avg. % Suburban:	64.6737
Avg. % Wilderness:	33.9757
Avg. Ratio of Rural to Suburban (Log):	-5.1815
Age of Restoration:	1990
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization Budget Class:	\$5m-\$150m
Ownership Type (PSCC):	County

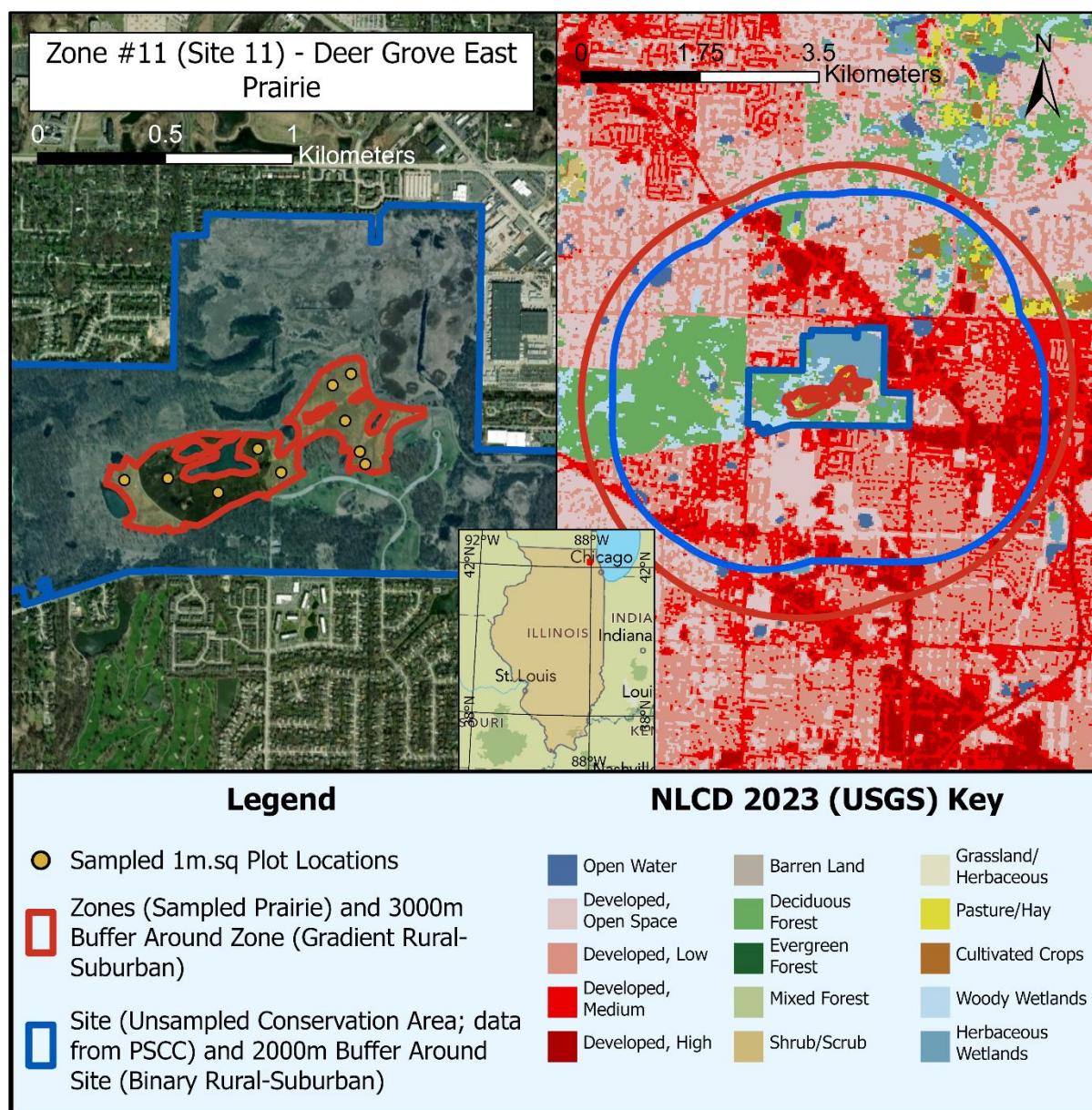
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Scirpus cyperinus*, *Sorghastrum nutans*, *Silphium integrifolium*, *Rosa carolina*, *Pycnanthemum virginianum*,

Top Five Highest Frequency/Coverage Native Species: *Solidago altissima*, *Monarda fistulosa*, *Packera paupercula*, *Andropogon gerardii*, *Silphium integrifolium*,

Top Five Highest Frequency/Coverage All Species: *Bromus inermis*, *Solidago altissima*, *Monarda fistulosa*, *Packera paupercula*, *Andropogon gerardii*,

Zone 11 (Site 11): Deer Grove East Prairie



Approximate Coordinates: 42°08'34"N 88°02'49"W

Date Sampled: September 2-3, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Fully Restoration

Operating Organization: Forest Preserves of Cook County

Full Species List: *Agrostis gigantea*, *Andropogon gerardii*, *Anemone canadensis*, *Baptisia alba*, *Carex scoparia*, *Coreopsis palmata*, *Coreopsis tripteris*, *Dalea purpurea*, *Desmodium illinoense*, *Dichanthelium oligosanthes*, *Drymocallis arguta*, *Echinacea pallida*, *Elymus canadensis*, *Epilobium leptophyllum*, *Eryngium yuccifolium*, *Euthamia graminifolia*, *Fragaria virginiana*, *Hackelia virginiana*, *Helianthus tuberosus*, *Heliopsis helianthoides*, *Hypericum perforatum*, *Lespedeza capitata*, *Liatris aspera*, *Lotus corniculatus*, *Monarda fistulosa*, *Oligoneuron rigidum*, *Parthenium integrifolium*, *Ratibida pinnata*, *Rhamnus cathartica*, *Rubus allegheniensis*, *Rubus occidentalis*, *Rudbeckia subtomentosa*, *Rudbeckia triloba*, *Schizachyrium scoparium*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium perfoliatum*, *Silphium terebinthinaceum*, *Solidago altissima*, *Solidago gigantea*, *Solidago juncea*, *Solidago nemoralis*, *Sorghastrum nutans*, *Spartina pectinata*, *Sporobolus heterolepis*, *Symphyotrichum ericoides*, *Symphyotrichum novae-angliae*, *Taraxacum officinale*, *Thalictrum dioicum*, *Trifolium hybridum*, *Trifolium pratense*, *Unidentified Asteraceae sp.*, *Zizia aptera*, *Zizia aurea*



Native Species Richness:	13.300
All Species Richness:	15.200
Introduced Species Percent:	12.706
Mean C (Native):	5.169
Mean C (All):	4.527
FQI (Native):	18.743
FQI (All):	16.566
AFQI:	48.173
Shannon Diversity (All):	3.570
Simpson Diversity (All):	0.963
Native Gamma Diversity:	46
All Gamma Diversity:	54

Zone Area (ha):	26.4513
Site Area (with Zone):	254.189
Distance to Chicago (km):	45.345
Avg. % Rural (8 Buffers):	0.46308
Avg. % Suburban:	53.9982
Avg. % Wilderness:	41.9158
Avg. Ratio of Rural to Suburban (Log):	-4.2466
Age of Restoration:	2009
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	> \$150m
Ownership Type (PSCC):	County

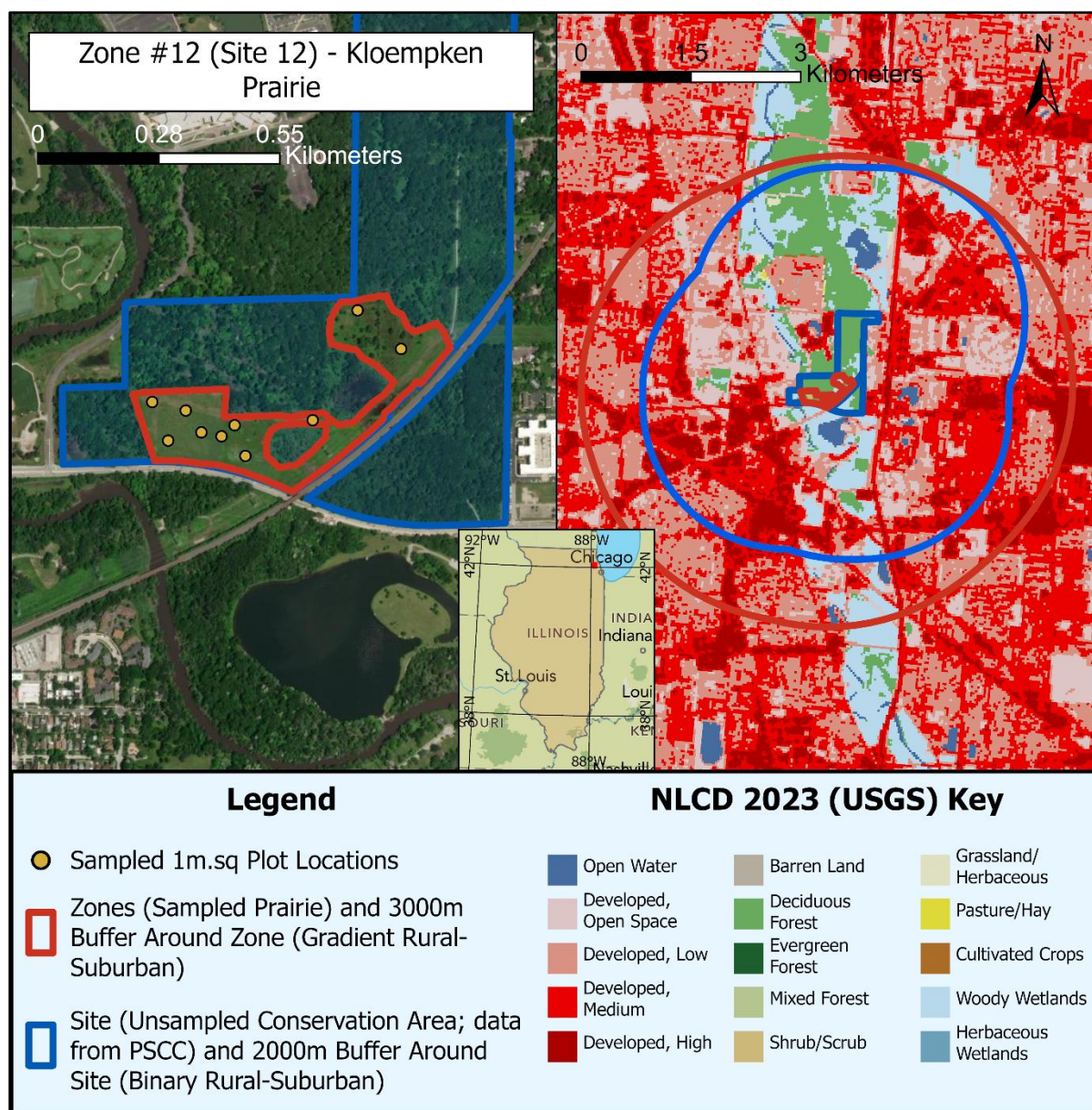
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Zizia aptera*, *Sporobolus heterolepsis*, *Epilobium leptophyllum*, *Echinacea pallida*, *Drymocallis arguta*,

Top Five Highest Frequency/Coverage Native Species: *Sorghastrum nutans*, *Zizia aurea*, *Parthenium integrifolium*, *Ratibida pinnata*, *Solidago altissima*,

Top Five Highest Frequency/Coverage All Species: *Sorghastrum nutans*, *Zizia aurea*, *Parthenium integrifolium*, *Ratibida pinnata*, *Solidago altissima*,

Zone 12 (Site 12): Kloempken Prairie



Approximate Coordinates: 42°03'19"N 87°52'50"W

Date Sampled: September 3-4, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Mixed

Operating Organization: Forest Preserves of Cook County

Full Species List: *Agrimonia parviflora*, *Andropogon gerardii*, *Baptisia alba*, *Bromus inermis*, *Cornus racemosa*, *Epilobium coloratum*, *Eryngium yuccifolium*, *Eupatorium altissimum*, *Eupatorium serotinum*, *Euthamia graminifolia*, *Fraxinus pennsylvanica*, *Helianthus grosseserratus*, *Hypericum perforatum*, *Lotus corniculatus*, *Lycopus americanus*, *Lycopus uniflorus*, *Monarda fistulosa*, *Penstemon digitalis*, *Pycnanthemum virginianum*, *Silphium integrifolium*, *Silphium laciniatum*, *Solidago altissima*, *Sorghastrum nutans*, *Ulmus americana*, *Vernonia missurica*



Native Species Richness:	4.300
All Species Richness:	5.200
Introduced Species Percent:	21.600
Mean C (Native):	4.238
Mean C (All):	3.292
FQI (Native):	8.262
FQI (All):	6.664
AFQI:	37.254
Shannon Diversity (All):	2.616
Simpson Diversity (All):	0.880
Native Gamma Diversity:	22
All Gamma Diversity:	25

Zone Area (ha):	9.63167
Site Area (with Zone):	71.5937
Distance to Chicago (km):	28.9949
Avg. % Rural (8 Buffers):	0.03855
Avg. % Suburban:	60.5535
Avg. % Wilderness:	36.9818
Avg. Ratio of Rural to Suburban (Log):	-7.638
Age of Restoration:	1988
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization	
Budget Class:	> \$150m
Ownership Type (PSCC):	County

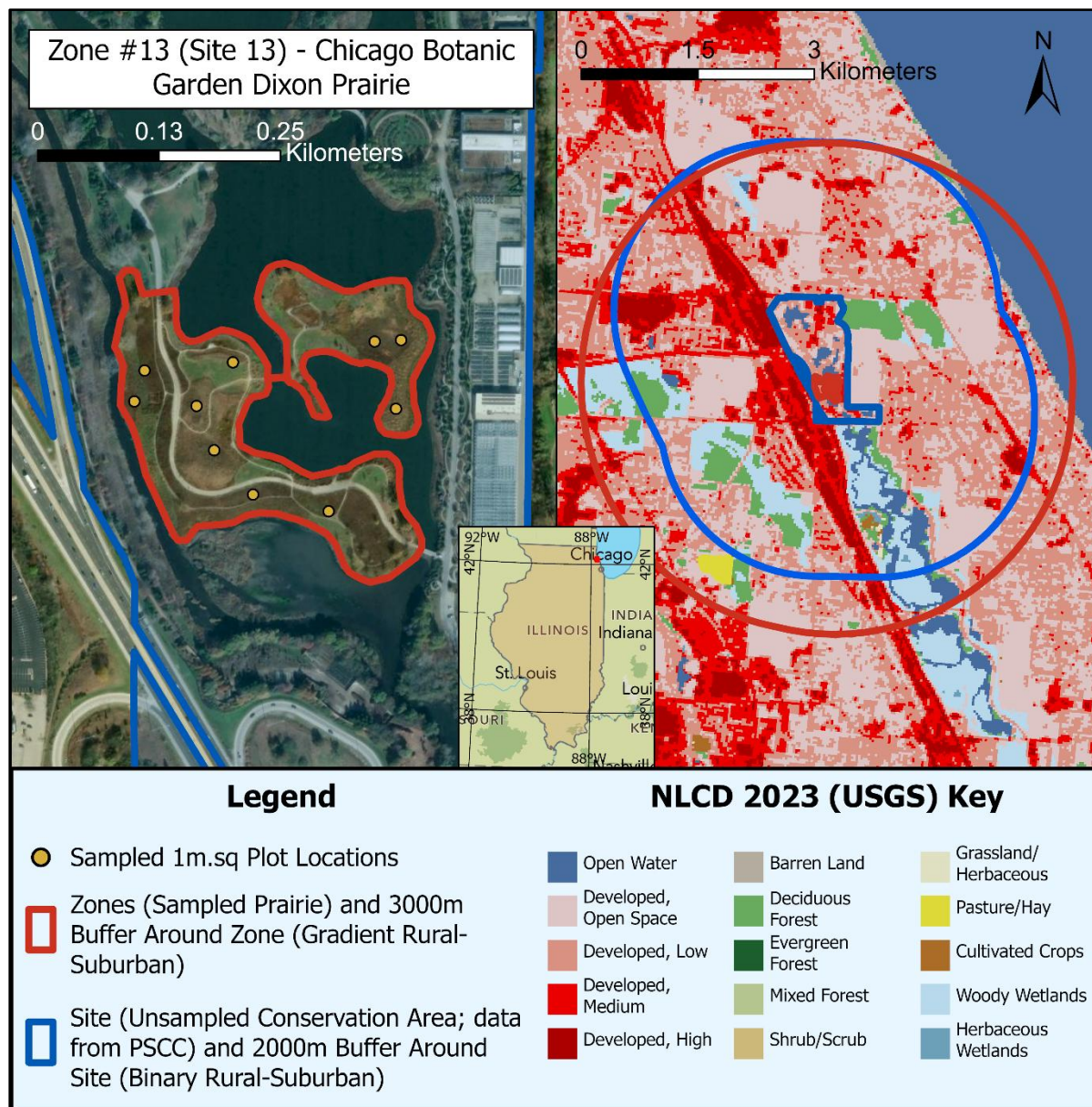
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Eryngium yuccifolium*, *Baptisia alba*, *Sorghastrum nutans*, *Silphium laciniatum*, *Silphium integrifolium*,

Top Five Highest Frequency/Coverage Native Species: *Andropogon gerardii*, *Euthamia graminifolia*, *Helianthus grosseserratus*, *Solidago altissima*, *Lycopus americanus*,

Top Five Highest Frequency/Coverage All Species: *Andropogon gerardii*, *Bromus inermis*, *Euthamia graminifolia*, *Helianthus grosseserratus*, *Solidago altissima*,

Zone 13 (Site 13): Chicago Botanic Garden Dixon Prairie



Approximate Coordinates: 42°08'33"N 87°47'13"W

Date Sampled: September 4, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Fully Restoration

Operating Organization: Chicago Botanic Garden

Full Species List: *Achillea millefolium*, *Amorpha canescens*, *Andropogon gerardii*, *Anemone canadensis*, *Apocynum cannabinum*, *Brickellia eupatorioides*, *Carex pellita*, *Coreopsis tripteris*, *Desmodium illinoense*, *Echinacea purpurea*, *Eryngium yuccifolium*, *Eupatorium altissimum*, *Euphorbia corollata*, *Euthamia graminifolia*, *Gaura biennis*, *Gentiana andrewsii*, *Geum canadense*, *Helianthus grosseserratus*, *Lespedeza capitata*, *Liatris aspera*, *Monarda fistulosa*, *Monarda punctata*, *Napaea dioica*, *Oligoneuron riddellii*, *Packera paupercula*, *Penstemon digitalis*, *Perideridia americana*, *Pycnanthemum tenuifolium*, *Pycnanthemum virginianum*, *Rosa carolina*, *Rudbeckia subtomentosa*, *Salix interior*, *Securigera varia*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium terebinthinaceum*, *Smilacina stellata*, *Solidago altissima*, *Solidago speciosa*, *Sporobolus heterolepis*, *Symphyotrichum ericoides*, *Symphyotrichum firmum*, *Symphyotrichum novae-angliae*, *Symphyotrichum pilosum*, *Symphyotrichum praealtum*, *Trifolium hybridum*, *Veronicastrum virginicum*



Native Species Richness:	9.600
All Species Richness:	10.100
Introduced Species Percent:	4.644
Mean C (Native):	4.969
Mean C (All):	4.761
FQI (Native):	15.189
FQI (All):	14.549
AFQI:	48.612
Shannon Diversity (All):	3.522
Simpson Diversity (All):	0.962
Native Gamma Diversity:	44
All Gamma Diversity:	47

Zone Area (ha):	5.64614
Site Area (with Zone):	113.474
Distance to Chicago (km):	34.8237
Avg. % Rural (8 Buffers):	0.05028
Avg. % Suburban:	80.0278
Avg. % Wilderness:	9.09957
Avg. Ratio of Rural to Suburban (Log):	-7.38
Age of Restoration:	1987
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization Budget Class:	\$5m-\$150m
Ownership Type (PSCC):	County

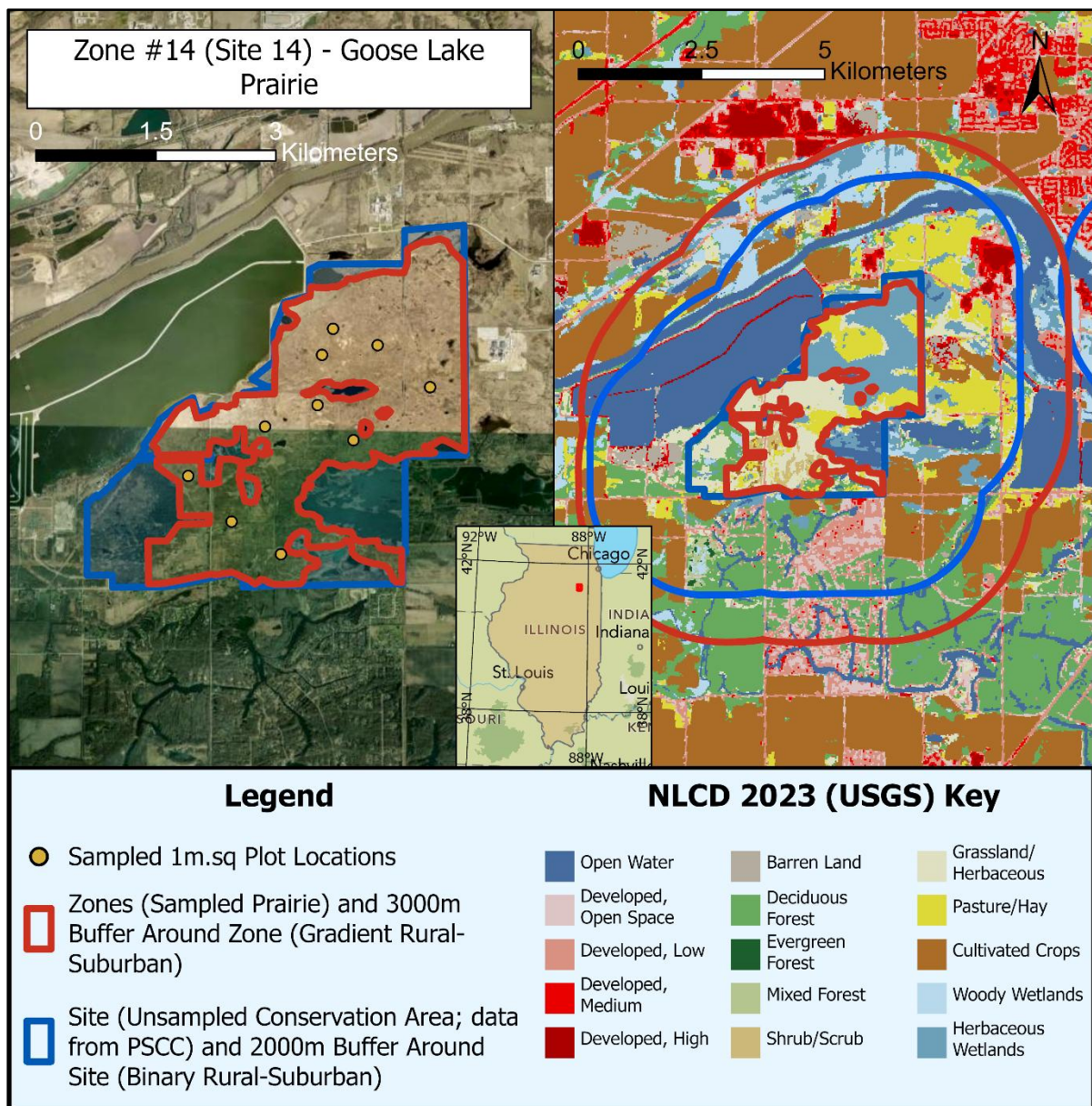
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Echinacea purpurea*, *Amorpha canescens*, *Gentiana andrewsii*, *Eryngium yuccifolium*,

Top Five Highest Frequency/Coverage Native Species: *Solidago altissima*, *Rosa carolina*, *Monarda fistulosa*, *Symphyotrichum ericoides*, *Euphorbia corollata*,

Top Five Highest Frequency/Coverage All Species: *Solidago altissima*, *Rosa carolina*, *Monarda fistulosa*, *Symphyotrichum ericoides*, *Euphorbia corollata*,

Zone 14 (Site 14): Goose Lake Prairie



Approximate Coordinates: 41°21'48"N 88°18'42"W

Date Sampled: September 5-6, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Remnant

Operating Organization: Illinois Department of Natural Resources

Full Species List: *Agrimonia parviflora*, *Agrostis gigantea*, *Ambrosia artemisiifolia*, *Andropogon gerardii*, *Apocynum cannabinum*, *Baptisia alba*, *Bromus inermis*, *Calystegia sepium*, *Chamaecrista fasciculata*, *Coreopsis tripteris*, *Dichanthelium oligosanthes*, *Elymus canadensis*, *Eryngium yuccifolium*, *Eupatorium altissimum*, *Euthamia graminifolia*, *Galium obtusum*, *Helianthus grosseserratus*, *Helianthus mollis*, *Lespedeza cuneata*, *Liatris pycnostachya*, *Monarda fistulosa*, *Parthenium integrifolium*, *Phlox glaberrima*, *Potentilla simplex*, *Pycnanthemum tenuifolium*, *Rosa carolina*, *Rubus allegheniensis*, *Rubus occidentalis*, *Rudbeckia subtomentosa*, *Securigera varia*, *Setaria faberi*, *Silphium integrifolium*, *Solidago altissima*, *Solidago gigantea*, *Solidago nemoralis*, *Solidago speciosa*, *Sorghastrum nutans*, *Sporobolus heterolepsis*, *Stachys pilosa*, *Symphotrichum ericoides*, *Vernonia missurica*



Native Species Richness:	7.600
All Species Richness:	8.400
Introduced Species Percent:	14.802
Mean C (Native):	4.072
Mean C (All):	3.562
FQI (Native):	11.384
FQI (All):	10.333
AFQI:	37.984
Shannon Diversity (All):	3.290
Simpson Diversity (All):	0.949
Native Gamma Diversity:	36
All Gamma Diversity:	41

Zone Area (ha):	875.725
Site Area (with Zone):	1297.64
Distance to Chicago (km):	73.8584
Avg. % Rural (8 Buffers):	11.1058
Avg. % Suburban:	12.8272
Avg. % Wilderness:	45.8406
Avg. Ratio of Rural to Suburban (Log):	-0.1972
Age of Restoration:	2000
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	> \$150m
Ownership Type (PSCC):	State

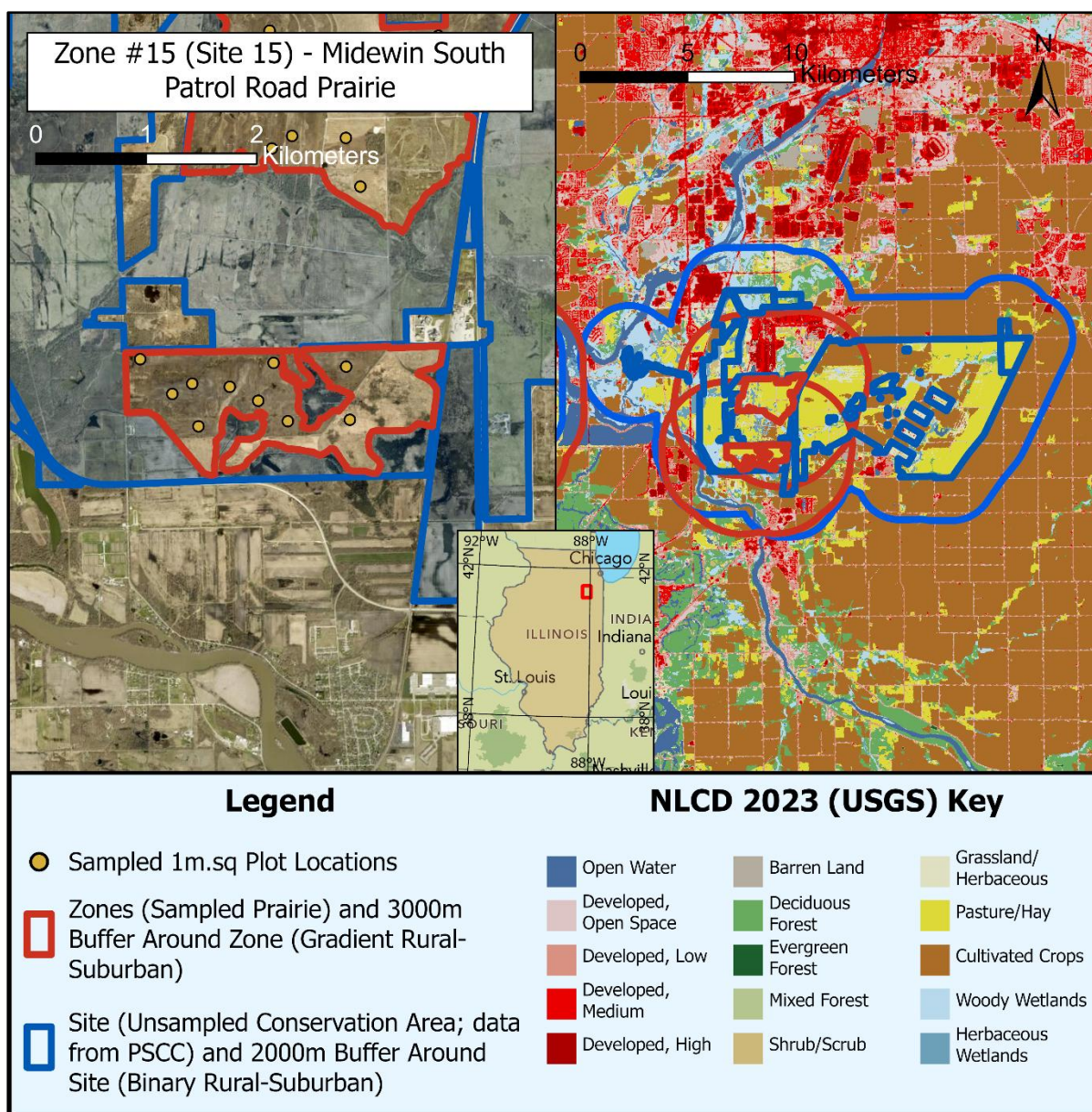
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Phlox glaberrima*, *Helianthus mollis*, *Eryngium yuccifolium*, *Dichanthelium oligosanthes*,

Top Five Highest Frequency/Coverage Native Species: *Solidago altissima*, *Sorghastrum nutans*, *Rubus occidentalis*, *Andropogon gerardii*, *Helianthus grosseserratus*,

Top Five Highest Frequency/Coverage All Species: *Solidago altissima*, *Sorghastrum nutans*, *Rubus occidentalis*, *Andropogon gerardii*, *Bromus inermis*,

Zone 15 (Site 15 – Midewin): Midewin South Patrol Road Prairie



Approximate Coordinates: 41°20'41"N 88°09'39"W

Date Sampled: September 6, 9, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Undefined

Operating Organization: U.S. Forest Service

Full Species List: *Amphicarpaea bracteata*, *Andropogon gerardii*, *Anemone canadensis*, *Apocynum cannabinum*, *Baptisia alba*, *Bromus inermis*, *Calystegia sepium*, *Chamaecrista fasciculata*, *Coreopsis palmata*, *Coreopsis tripteris*, *Desmanthus illinoensis*, *Desmodium canadense*, *Elymus canadensis*, *Eryngium yuccifolium*, *Euthamia graminifolia*, *Fragaria virginiana*, *Helianthus decapetalus*, *Heliopsis helianthoides*, *Monarda fistulosa*, *Oligoneuron rigidum*, *Parthenium integrifolium*, *Penstemon digitalis*, *Physostegia virginiana*, *Ratibida pinnata*, *Rudbeckia subtomentosa*, *Securigera varia*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium perfoliatum*, *Silphium terebinthinaceum*, *Solidago altissima*, *Solidago canadensis*, *Solidago gigantea*, *Solidago nemoralis*, *Sorghastrum nutans*, *Symphytotrichum novae-angliae*, *Symphytotrichum pilosum*, *Thalictrum dioicum*, *Trifolium hybridum*, *Trifolium pratense*, *Zizia aurea*



Native Species Richness:	12.200
All Species Richness:	13.100
Introduced Species Percent:	6.251
Mean C (Native):	4.711
Mean C (All):	4.415
FQI (Native):	16.219
FQI (All):	15.169
AFQI:	45.584
Shannon Diversity (All):	3.301
Simpson Diversity (All):	0.950
Native Gamma Diversity:	37
All Gamma Diversity:	41

Zone Area (ha):	215.585
Site Area (with Zone):	7245.6
Distance to Chicago (km):	67.2311
Avg. % Rural (8 Buffers):	15.2848
Avg. % Suburban:	15.0345
Avg. % Wilderness:	39.1699
Avg. Ratio of Rural to Suburban (Log):	0.10874
Age of Restoration:	Undefined
Sector Type (PSCC):	Public
Special Legal Designations:	Undefined
Operating Organization	
Budget Class:	Undefined
Ownership Type (PSCC):	Federal

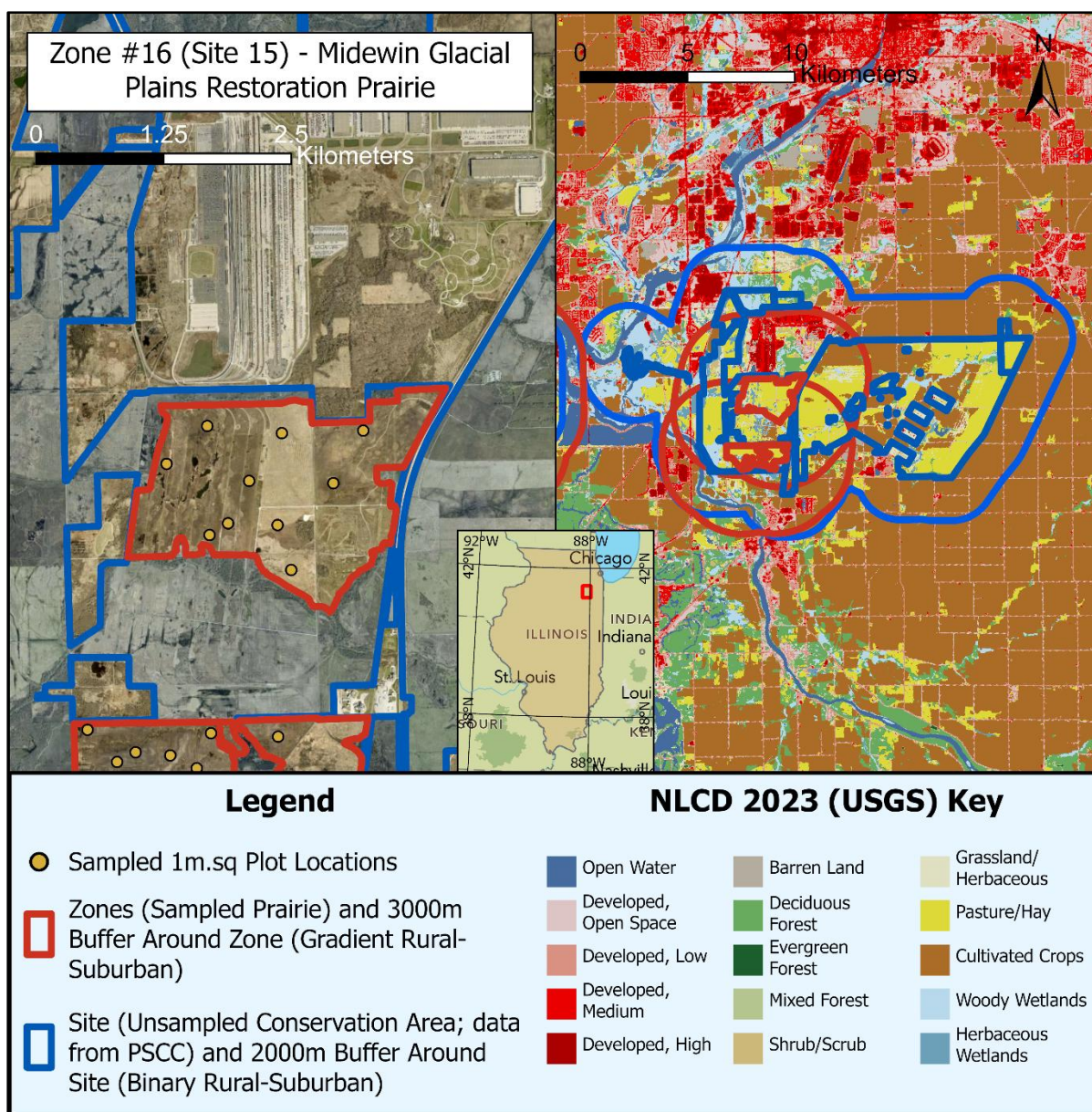
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Coreopsis palmata*, *Eryngium yuccifolium*, *Rudbeckia subtomentosa*, *Parthenium integrifolium*, *Helianthus decapetalus*,

Top Five Highest Frequency/Coverage Native Species: *Solidago altissima*, *Andropogon gerardii*, *Zizia aurea*, *Desmodium canadense*, *Rudbeckia subtomentosa*,

Top Five Highest Frequency/Coverage All Species: *Solidago altissima*, *Andropogon gerardii*, *Zizia aurea*, *Desmodium canadense*, *Rudbeckia subtomentosa*,

Zone 16 (Site 15 - Midewin): Midewin Glacial Plains Prairie



Approximate Coordinates: 41°22'01"N 88°09'09"W

Date Sampled: September 9, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Undefined

Operating Organization: U.S. Forest Service

Notes of Interest: Previously a munitions storage depot for a U.S. military base.

Full Species List: *Achillea millefolium*, *Agrostis gigantea*, *Ambrosia artemisiifolia*, *Asclepias syriaca*, *Bouteloua curtipendula*, *Bromus inermis*, *Calystegia sepium*, *Coreopsis tripteris*, *Daucus carota*, *Echinacea pallida*, *Eryngium yuccifolium*, *Eupatorium altissimum*, *Euthamia graminifolia*, *Fragaria virginiana*, *Helianthus grosseserratus*, *Heliopsis helianthoides*, *Juncus greenii*, *Leucanthemum vulgare*, *Monarda fistulosa*, *Morus alba*, *Oligoneuron rigidum*, *Panicum virgatum*, *Penstemon digitalis*, *Perideridia americana*, *Phleum pratense*, *Phlox paniculata*, *Phryma leptostachya*, *Potentilla recta*, *Pycnanthemum tenuifolium*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rubus occidentalis*, *Rudbeckia subtomentosa*, *Schizachyrium scoparium*, *Silphium integrifolium*, *Silphium laciniatum*, *Solanum carolinense*, *Solidago altissima*, *Solidago speciosa*, *Symphotrichum ericoides*, *Symphotrichum lateriflorum*, *Symphotrichum pilosum*, *Trifolium hybridum*, *Trifolium pratense*, *Unidentified sp.*



Native Species Richness:	8.000
All Species Richness:	10.800
Introduced Species Percent:	30.659
Mean C (Native):	3.226
Mean C (All):	2.461
FQI (Native):	9.862
FQI (All):	7.791
AFQI:	27.900
Shannon Diversity (All):	3.297
Simpson Diversity (All):	0.940
Native Gamma Diversity:	34
All Gamma Diversity:	45

Zone Area (ha):	398.506
Site Area (with Zone):	7245.6
Distance to Chicago (km):	64.6899
Avg. % Rural (8 Buffers):	5.19013
Avg. % Suburban:	17.8133
Avg. % Wilderness:	35.1025
Avg. Ratio of Rural to Suburban (Log):	-1.1971
Age of Restoration:	Undefined
Sector Type (PSCC):	Public
Special Legal Designations:	Undefined
Operating Organization	
Budget Class:	Undefined
Ownership Type (PSCC):	Federal

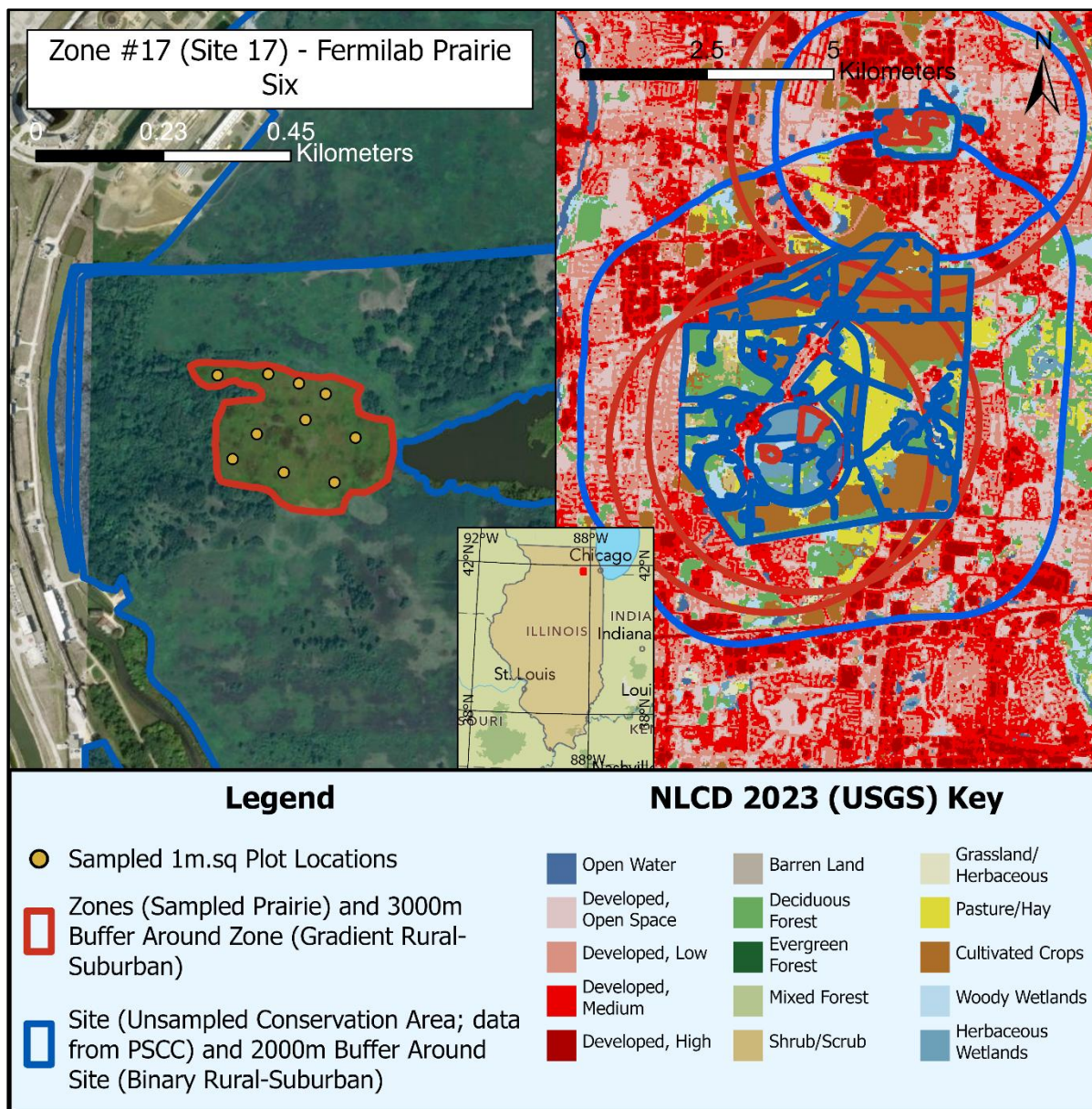
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Echinacea pallida*, *Eryngium yuccifolium*, *Solidago speciosa*, *Rudbeckia subtomentosa*, *Perideridia americana*,

Top Five Highest Frequency/Coverage Native Species: *Solidago altissima*, *Rudbeckia subtomentosa*, *Monarda fistulosa*, *Panicum virgatum*, *Euthamia graminifolia*,

Top Five Highest Frequency/Coverage All Species: *Solidago altissima*, *Bromus inermis*, *Trifolium pratense*, *Agrostis gigantea*, *Rudbeckia subtomentosa*,

Zone 17 (Site 17 - Fermilab): Fermilab Prairie Six



Approximate Coordinates: 41°49'52"N 88°15'27"W

Date Sampled: September 10, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Fully Restoration

Operating Organization: U.S. Department of Energy

Notes of Interest: Fermilab is home to four particle accelerators and a neutrino lab.

Full Species List: *Agrimonia gryposepala*, *Agrostis gigantea*, *Amorpha canescens*, *Andropogon gerardii*, *Anemone cylindrica*, *Apocynum cannabinum*, *Asclepias syriaca*, *Baptisia alba*, *Comandra umbellata*, *Coreopsis palmata*, *Coreopsis tripteris*, *Cornus racemosa*, *Dalea candida*, *Dalea purpurea*, *Desmodium canadense*, *Dichanthelium oligosanthes*, *Echinacea pallida*, *Equisetum arvense*, *Eryngium yuccifolium*, *Eupatorium altissimum*, *Euphorbia corollata*, *Euthamia graminifolia*, *Gentiana alba*, *Gentiana andrewsii*, *Gentianella quinquefolia*, *Helianthus grosseserratus*, *Helianthus pauciflorus*, *Hypericum perforatum*, *Lespedeza capitata*, *Lycopus americanus*, *Melilotus alba*, *Monarda fistulosa*, *Parthenium integrifolium*, *Pedicularis canadensis*, *Penstemon digitalis*, *Phlox pilosa*, *Physostegia virginiana*, *Pycnanthemum virginianum*, *Rudbeckia subtomentosa*, *Securigera varia*, *Silphium integrifolium*, *Silphium terebinthinaceum*, *Solidago altissima*, *Sorghastrum nutans*, *Sporobolus heterolepis*, *Symphyotrichum ericoides*, *Unidentified Cuscuta sp.*, *Veronicastrum virginicum*, *Viola sororia*, *Zizia aurea*



Native Species Richness:	14.300
All Species Richness:	15.400
Introduced Species Percent:	5.657
Mean C (Native):	4.878
Mean C (All):	4.604
FQI (Native):	18.584
FQI (All):	17.603
AFQI:	46.945
Shannon Diversity (All):	3.482
Simpson Diversity (All):	0.953
Native Gamma Diversity:	45
All Gamma Diversity:	50

Zone Area (ha):	6.23043
Site Area (with Zone):	2226.81
Distance to Chicago (km):	47.4693
Avg. % Rural (8 Buffers):	6.502
Avg. % Suburban:	26.2036
Avg. % Wilderness:	59.487
Avg. Ratio of Rural to Suburban (Log):	-1.5968
Age of Restoration:	1979
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization	
Budget Class:	> \$150m
Ownership Type (PSCC):	Federal

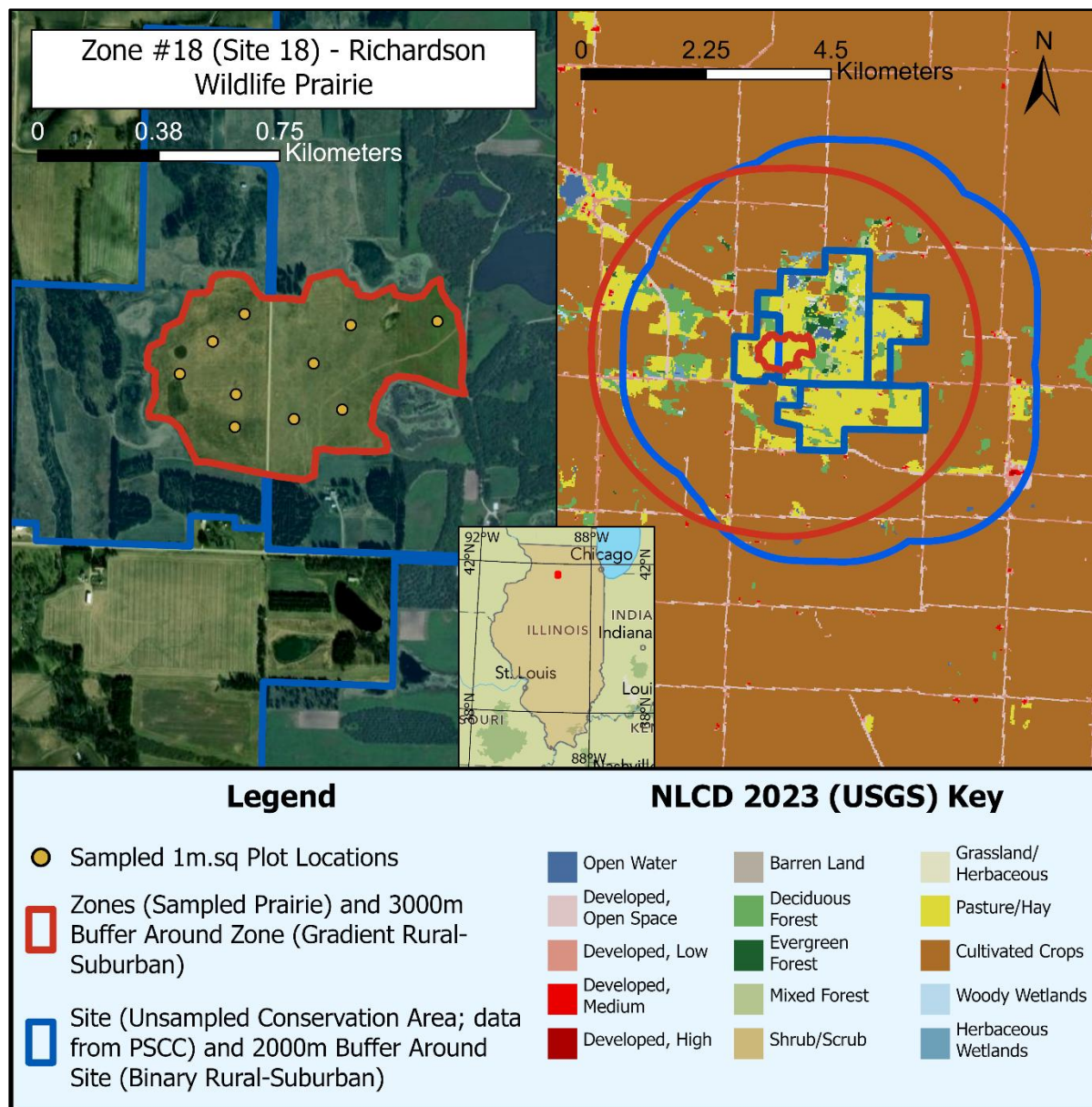
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Echinacea pallida*, *Dalea candida*, *Coreopsis palmata*, *Amorpha canescens*,

Top Five Highest Frequency/Coverage Native Species: *Andropogon gerardii*, *Monarda fistulosa*, *Sorghastrum nutans*, *Cornus racemosa*, *Solidago altissima*,

Top Five Highest Frequency/Coverage All Species: *Andropogon gerardii*, *Monarda fistulosa*, *Sorghastrum nutans*, *Cornus racemosa*, *Solidago altissima*,

Zone 18 (Site 18): Richardson Wildlife Prairie



Approximate Coordinates: 41°42'45"N 89°11'48"W

Date Sampled: September 11, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Restoration

Operating Organization: Richardson Wildlife Foundation

Full Species List: *Andropogon gerardii*, *Apocynum cannabinum*, *Asclepias syriaca*, *Baptisia alba*, *Botrypus virginianus*, *Coreopsis tripteris*, *Desmodium canadense*, *Echinacea purpurea*, *Elymus canadensis*, *Eryngium yuccifolium*, *Euthamia graminifolia*, *Gentiana alba*, *Helianthus grosseserratus*, *Heliopsis helianthoides*, *Lespedeza capitata*, *Melilotus alba*, *Monarda fistulosa*, *Morus alba*, *Oligoneuron rigidum*, *Parthenium integrifolium*, *Penstemon digitalis*, *Ratibida pinnata*, *Rubus occidentalis*, *Rudbeckia subtomentosa*, *Silphium integrifolium*, *Silphium perfoliatum*, *Solidago altissima*, *Solidago gigantea*, *Sorghastrum nutans*, *Sporobolus heterolepis*, *Stachys pilosa*, *Symphyotrichum pilosum*, *Taraxacum officinale*, *Thalictrum dasycarpum*, *Trifolium pratense*, *Zizia aurea*



Native Species Richness:	12.100
All Species Richness:	12.700
Introduced Species Percent:	4.577
Mean C (Native):	4.881
Mean C (All):	4.665
FQI (Native):	17.002
FQI (All):	16.257
AFQI:	47.701
Shannon Diversity (All):	3.141
Simpson Diversity (All):	0.940
Native Gamma Diversity:	32
All Gamma Diversity:	36

Zone Area (ha):	42.8132
Site Area (with Zone):	783.383
Distance to Chicago (km):	126.462
Avg. % Rural (8 Buffers):	33.3621
Avg. % Suburban:	5.13989
Avg. % Wilderness:	16.0884
Avg. Ratio of Rural to Suburban (Log):	1.83513
Age of Restoration:	1996
Sector Type (PSCC):	Private
Special Legal Designations:	No
Operating Organization	
Budget Class:	< \$5m
Ownership Type (PSCC):	NGO

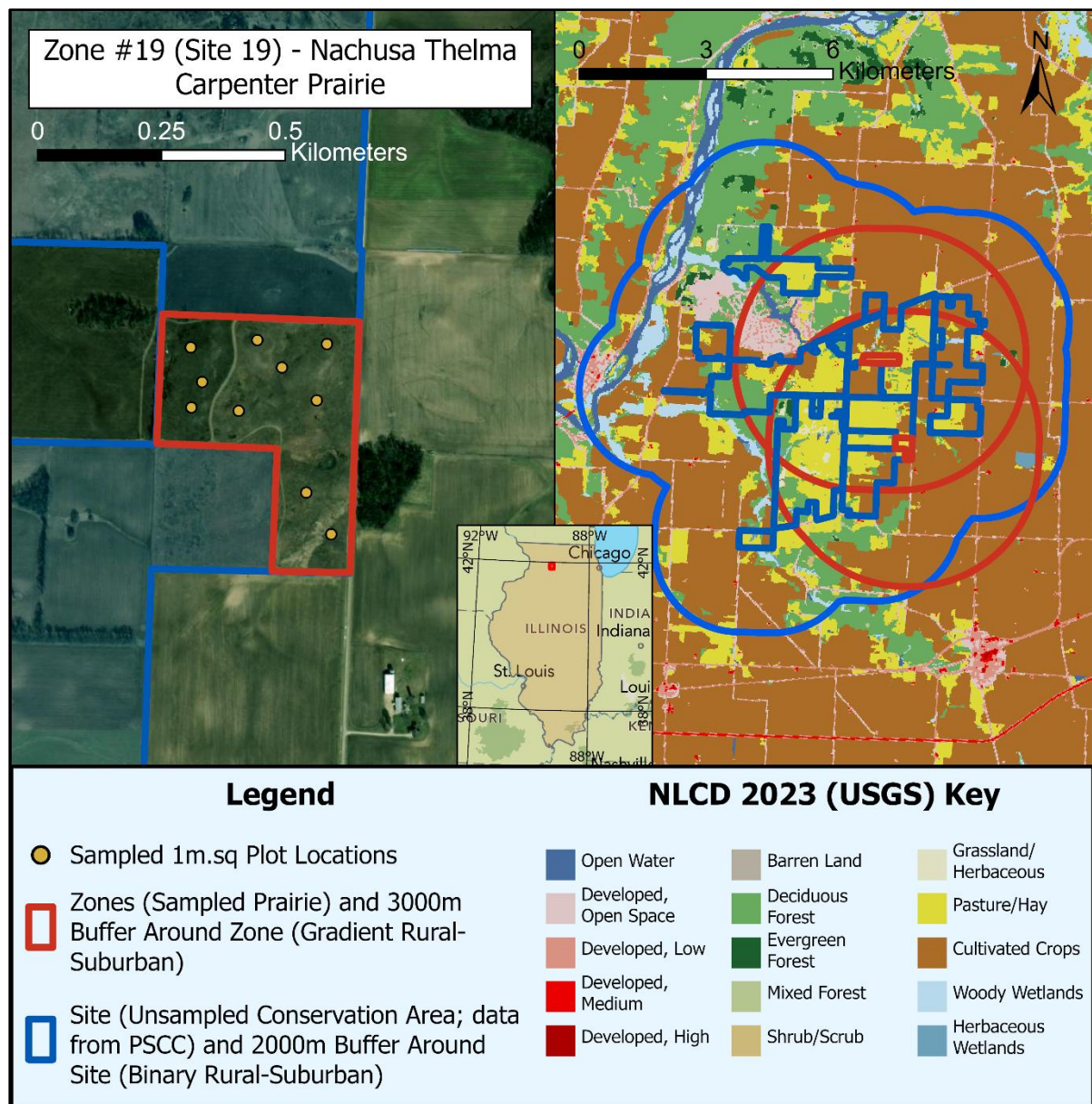
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Echinacea purpurea*, *Gentiana alba*, *Eryngium yuccifolium*, *Rudbeckia subtomentosa*,

Top Five Highest Frequency/Coverage Native Species: *Andropogon gerardii*, *Solidago altissima*, *Sorghastrum nutans*, *Baptisia alba*, *Coreopsis tripteris*,

Top Five Highest Frequency/Coverage All Species: *Andropogon gerardii*, *Solidago altissima*, *Sorghastrum nutans*, *Baptisia alba*, *Coreopsis tripteris*,

Zone 19 (Site 19 - Nachusa): Nachusa Thelma Carpenter Prairie



Approximate Coordinates: 41°53'07"N 89°19'24"W

Date Sampled: September 12, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Remnant

Operating Organization: The Nature Conservancy

Full Species List: *Achillea millefolium*, *Andropogon gerardii*, *Anemone cylindrica*, *Apocynum cannabinum*, *Asclepias syriaca*, *Asclepias verticillata*, *Bouteloua curtipendula*, *Bromus inermis*, *Comandra umbellata*, *Dalea purpurea*, *Daucus carota*, *Dichanthelium scribnerianum*, *Echinacea pallida*, *Eupatorium altissimum*, *Euthamia graminifolia*, *Gentiana puberulenta*, *Isanthus brachiatus*, *Liatris aspera*, *Medicago lupulina*, *Monarda fistulosa*, *Oligoneuron rigidum*, *Oxalis stricta*, *Pastinaca sativa*, *Phleum pratense*, *Ratibida pinnata*, *Rosa carolina*, *Rubus occidentalis*, *Schizachyrium scoparium*, *Scutellaria parvula*, *Solidago altissima*, *Solidago gigantea*, *Solidago nemoralis*, *Solidago speciosa*, *Sorghastrum nutans*, *Sporobolus heterolepis*, *Symphyotrichum ericoides*, *Symphyotrichum oblongifolium*, *Symphyotrichum pilosum*, *Trifolium pratense*, *Viola sororia*, *Vitis riparia*



Native Species Richness:	7.900
All Species Richness:	9.200
Introduced Species Percent:	21.084
Mean C (Native):	4.200
Mean C (All):	3.768
FQI (Native):	12.492
FQI (All):	11.299
AFQI:	39.722
Shannon Diversity (All):	3.378
Simpson Diversity (All):	0.957
Native Gamma Diversity:	35
All Gamma Diversity:	41

Zone Area (ha):	14.3206
Site Area (with Zone):	1885.92
Distance to Chicago (km):	136.05
Avg. % Rural (8 Buffers):	53.1127
Avg. % Suburban:	5.98194
Avg. % Wilderness:	6.35852
Avg. Ratio of Rural to Suburban (Log):	2.44303
Age of Restoration:	Remnant
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	< \$5m
Ownership Type (PSCC):	State

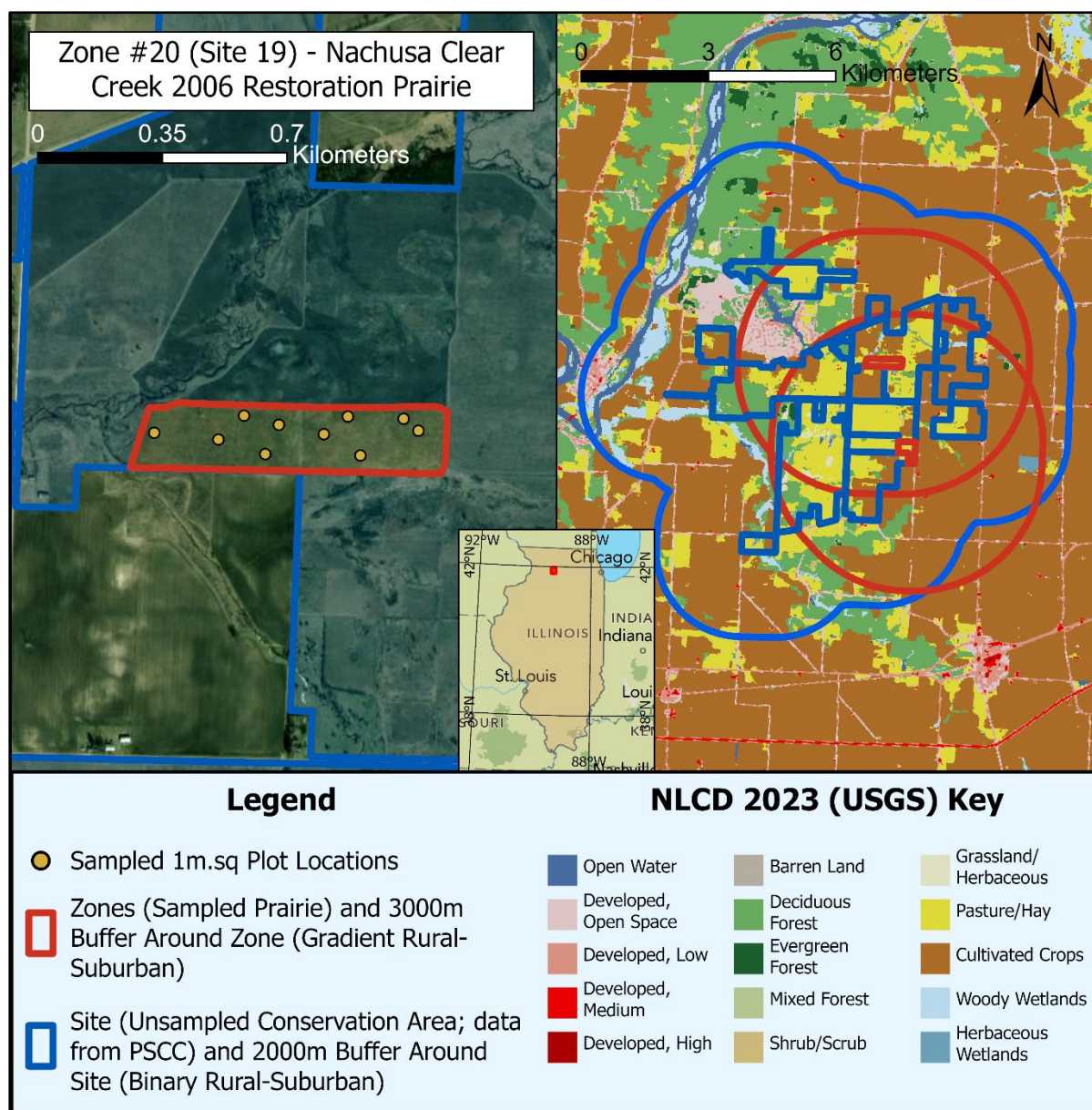
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Gentiana puberulenta*, *Echinacea pallida*, *Symphyotrichum oblongifolium*, *Dalea purpurea*,

Top Five Highest Frequency/Coverage Native Species: *Schizachyrium scoparium*, *Euthamia graminifolia*, *Sporobolus heterolepsis*, *Rubus occidentalis*, *Symphyotrichum ericoides*,

Top Five Highest Frequency/Coverage All Species: *Bromus inermis*, *Schizachyrium scoparium*, *Euthamia graminifolia*, *Sporobolus heterolepsis*, *Rubus occidentalis*,

Zone 20 (Site 19 - Nachusa): Nachusa Clear Creek 2006 Prairie



Approximate Coordinates: 41°54'12"N 89°19'53"W

Date Sampled: September 12-13, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Restoration

Operating Organization: The Nature Conservancy

Full Species List: *Andropogon gerardii*, *Anemone cylindrica*, *Astragalus canadensis*, *Baptisia alba*, *Bouteloua curtipendula*, *Brickellia eupatorioides*, *Coreopsis palmata*, *Coreopsis tripteris*, *Dalea purpurea*, *Desmodium canadense*, *Desmodium illinoense*, *Drymocallis arguta*, *Echinacea pallida*, *Elymus canadensis*, *Eryngium yuccifolium*, *Euthamia graminifolia*, *Gentiana alba*, *Helianthus pauciflorus*, *Lespedeza capitata*, *Liatris aspera*, *Monarda fistulosa*, *Morus alba*, *Oligoneuron rigidum*, *Parthenium integrifolium*, *Penstemon digitalis*, *Phleum pratense*, *Prunus serotina*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rudbeckia subtomentosa*, *Schizachyrium scoparium*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium terebinthinaceum*, *Solidago altissima*, *Solidago gigantea*, *Solidago speciosa*, *Sorghastrum nutans*, *Symphyotrichum ericoides*, *Symphyotrichum pilosum*, *Tephrosia virginiana*, *Trifolium pratense*, *Zizia aurea*



Native Species Richness:	16.600
All Species Richness:	17.700
Introduced Species Percent:	6.398
Mean C (Native):	5.940
Mean C (All):	5.562
FQI (Native):	24.198
FQI (All):	22.696
AFQI:	57.473
Shannon Diversity (All):	3.325
Simpson Diversity (All):	0.952
Native Gamma Diversity:	40
All Gamma Diversity:	43

Zone Area (ha):	15.5639
Site Area (with Zone):	1885.92
Distance to Chicago (km):	136.749
Avg. % Rural (8 Buffers):	38.6234
Avg. % Suburban:	5.4275
Avg. % Wilderness:	12.5584
Avg. Ratio of Rural to Suburban (Log):	2.03837
Age of Restoration:	2006
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	< \$5m
Ownership Type (PSCC):	State

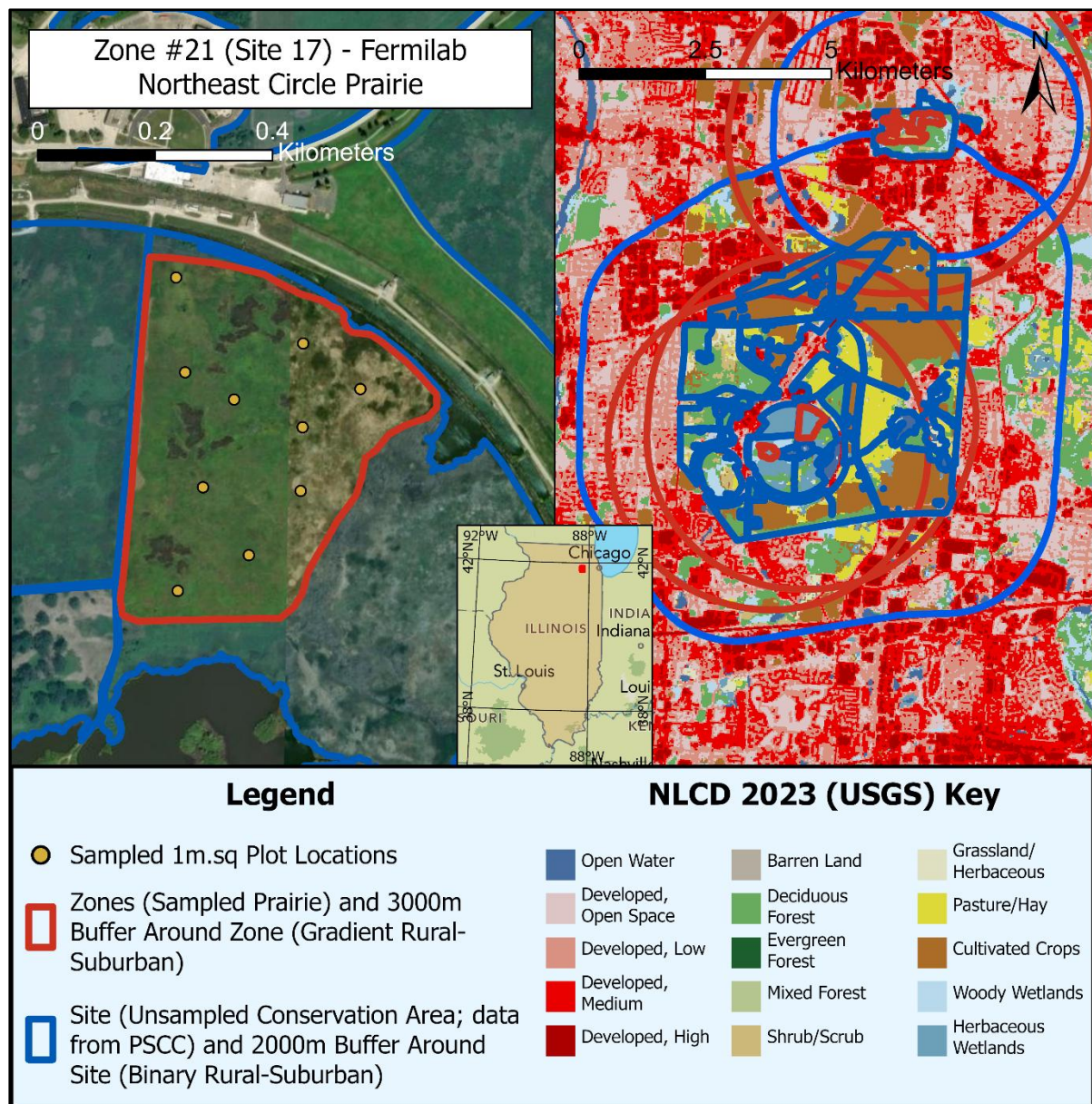
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Echinacea pallida*, *Drymocallis arguta*, *Coreopsis palmata*, *Helianthus pauciflorus*, *Gentiana alba*,

Top Five Highest Frequency/Coverage Native Species: *Sorghastrum nutans*, *Zizia aurea*, *Ratibida pinnata*, *Coreopsis palmata*, *Echinacea pallida*,

Top Five Highest Frequency/Coverage All Species: *Sorghastrum nutans*, *Zizia aurea*, *Ratibida pinnata*, *Coreopsis palmata*, *Echinacea pallida*,

Zone 21 (Site 17 - Fermilab): Fermilab Northeast Circle Prairie



Approximate Coordinates: 41°50'11"N 88°14'52"W

Date Sampled: September 16, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Fully Restoration

Operating Organization: U.S. Department of Energy

Full Species List: *Ambrosia artemisiifolia*, *Ambrosia trifida*, *Andropogon gerardii*, *Apocynum cannabinum*, *Baptisia alba*, *Bromus inermis*, *Cirsium arvense*, *Coreopsis tripteris*, *Cornus racemosa*, *Echinacea purpurea*, *Eupatorium altissimum*, *Euphorbia corollata*, *Euthamia graminifolia*, *Gentiana alba*, *Helianthus grosseserratus*, *Heliopsis helianthoides*, *Lespedeza capitata*, *Melilotus alba*, *Monarda fistulosa*, *Nabalus asper*, *Oligoneuron rigidum*, *Oxypolis rigidior*, *Parthenium integrifolium*, *Pedicularis canadensis*, *Penstemon digitalis*, *Phalaris arundinacea*, *Physostegia virginiana*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rubus occidentalis*, *Rudbeckia subtomentosa*, *Securigera varia*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium terebinthinaceum*, *Smilax lasioneuron*, *Solidago altissima*, *Solidago nemoralis*, *Sorghastrum nutans*, *Sporobolus heterolepsis*, *Symphyotrichum drummondii*, *Symphyotrichum ericoides*, *Veronicastrum virginicum*, *Viola sororia*, *Zizia aurea*



Native Species Richness:	9.100
All Species Richness:	10.600
Introduced Species Percent:	16.527
Mean C (Native):	4.126
Mean C (All):	3.530
FQI (Native):	12.741
FQI (All):	11.038
AFQI:	38.104
Shannon Diversity (All):	3.155
Simpson Diversity (All):	0.926
Native Gamma Diversity:	40
All Gamma Diversity:	45

Zone Area (ha):	22.4391
Site Area (with Zone):	2226.81
Distance to Chicago (km):	46.6694
Avg. % Rural (8 Buffers):	11.1066
Avg. % Suburban:	29.4394
Avg. % Wilderness:	43.7666
Avg. Ratio of Rural to Suburban (Log):	-0.9795
Age of Restoration:	1979
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization	
Budget Class:	> \$150m
Ownership Type (PSCC):	Federal

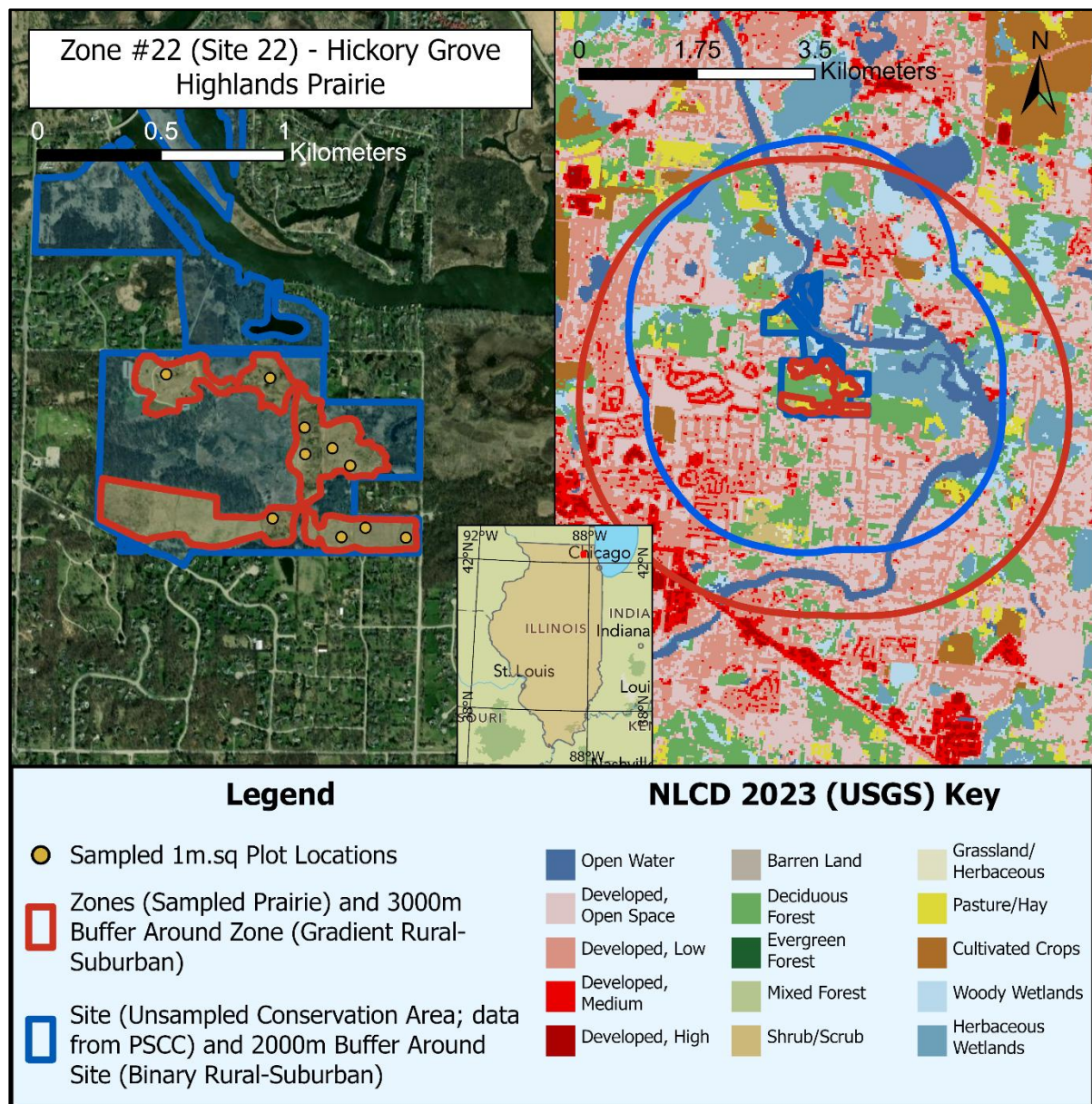
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Nabalus asper*, *Echinacea purpurea*, *Pedicularis canadensis*, *Gentiana alba*,

Top Five Highest Frequency/Coverage Native Species: *Solidago altissima*, *Cornus racemosa*, *Monarda fistulosa*, *Andropogon gerardii*, *Sorghastrum nutans*,

Top Five Highest Frequency/Coverage All Species: *Securigera varia*, *Solidago altissima*, *Cornus racemosa*, *Monarda fistulosa*, *Andropogon gerardii*,

Zone 22 (Site 22): Hickory Grove Highlands Prairie



Approximate Coordinates: 42°13'50"N 88°12'22"W

Date Sampled: September 17, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Fully Restoration

Operating Organization: McHenry County Conservation District

Full Species List: *Agrostis gigantea*, *Andropogon gerardii*, *Anemone virginiana*, *Apocynum cannabinum*, *Asclepias tuberosa*, *Bromus inermis*, *Cirsium altissimum*, *Coreopsis tripteris*, *Cornus racemosa*, *Daucus carota*, *Euthamia graminifolia*, *Frangula alnus*, *Gentiana alba*, *Melilotus alba*, *Monarda fistulosa*, *Parthenium integrifolium*, *Phalaris arundinacea*, *Pycnanthemum tenuifolium*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Silphium integrifolium*, *Silphium laciniatum*, *Solidago altissima*, *Solidago gigantea*, *Solidago juncea*, *Solidago nemoralis*, *Sorghastrum nutans*, *Stachys hispida*, *Symphyotrichum lateriflorum*, *Symphyotrichum pilosum*, *Taraxacum officinale*, *Trifolium pratense*, *Verbena urticifolia*, *Viola sororia*, *Vitis riparia*



Native Species Richness:	5.300
All Species Richness:	7.200
Introduced Species Percent:	24.841
Mean C (Native):	3.975
Mean C (All):	3.009
FQI (Native):	8.697
FQI (All):	6.557
AFQI:	34.491
Shannon Diversity (All):	3.037
Simpson Diversity (All):	0.930
Native Gamma Diversity:	27
All Gamma Diversity:	35

Zone Area (ha):	35.6881
Site Area (with Zone):	137.225
Distance to Chicago (km):	61.2978
Avg. % Rural (8 Buffers):	0.60304
Avg. % Suburban:	47.2123
Avg. % Wilderness:	43.8248
Avg. Ratio of Rural to Suburban (Log):	-4.4916
Age of Restoration:	2000
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization Budget Class:	\$5m-\$150m
Ownership Type (PSCC):	Municipal

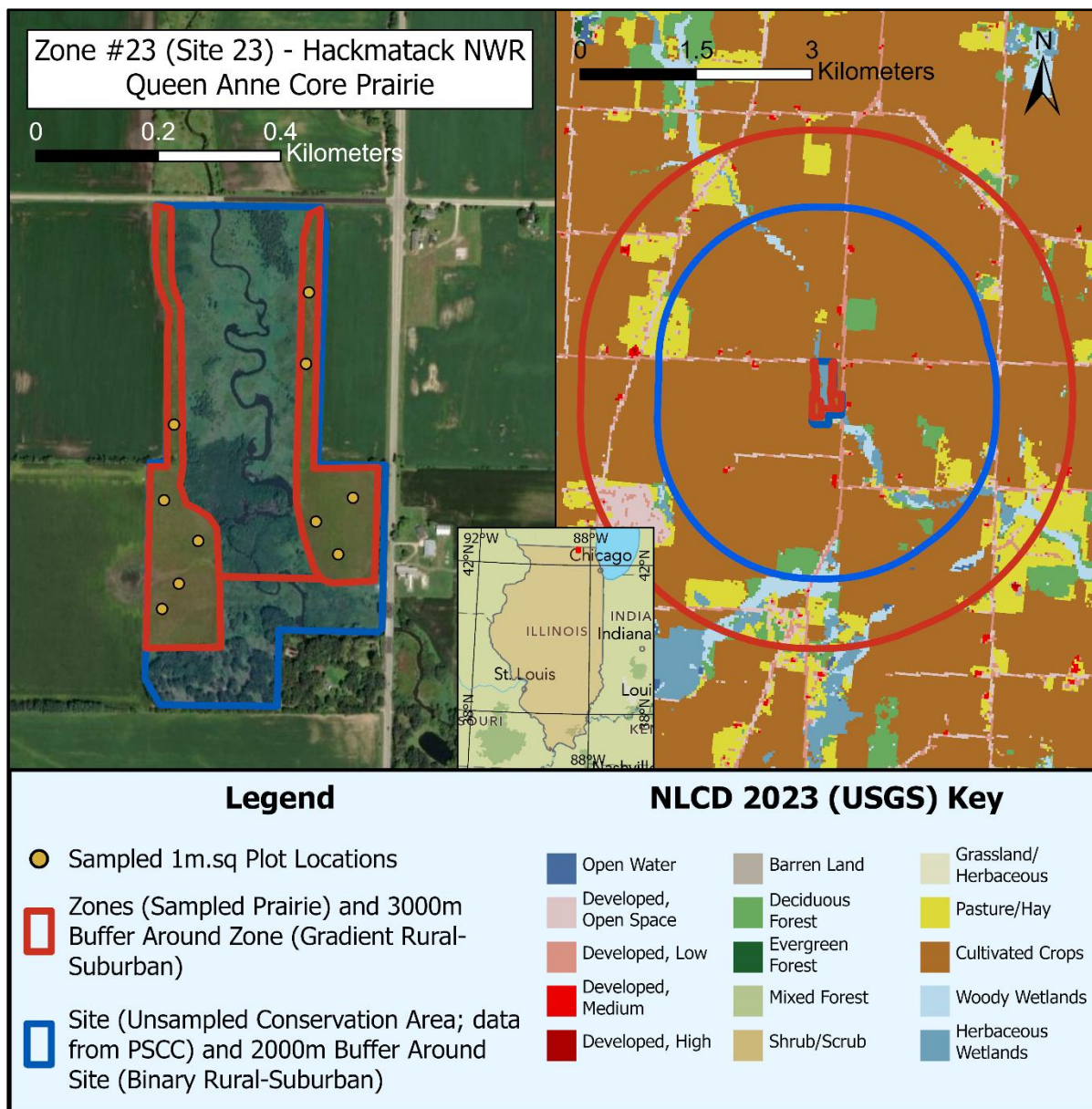
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Gentiana alba*, *Parthenium integrifolium*, *Cirsium altissimum*, *Asclepias tuberosa*, *Pycnanthemum tenuifolium*,

Top Five Highest Frequency/Coverage Native Species: *Sorghastrum nutans*, *Solidago altissima*, *Andropogon gerardii*, *Cornus racemosa*, *Solidago gigantea*,

Top Five Highest Frequency/Coverage All Species: *Sorghastrum nutans*, *Bromus inermis*, *Solidago altissima*, *Andropogon gerardii*, *Cornus racemosa*,

Zone 23 (Site 23): Hackmatack NWR Queen Anne Core Prairie



*NWR stands for National Wildlife Refuge

Approximate Coordinates: 42°24'34"N 88°26'21"W

Date Sampled: September 18, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Restoration

Operating Organization: U.S. Fish and Wildlife Service

Full Species List: *Agrostis gigantea*, *Bromus inermis*, *Carex cristatella*, *Cirsium arvense*, *Coreopsis tripteris*, *Daucus carota*, *Desmodium canadense*, *Drymocallis arguta*, *Elymus canadensis*, *Elymus virginicus*, *Eupatorium altissimum*, *Gaura biennis*, *Geum laciniatum*, *Helenium autumnale*, *Heliopsis helianthoides*, *Monarda fistulosa*, *Oligoneuron rigidum*, *Penstemon digitalis*, *Phalaris arundinacea*, *Plantago major*, *Prunella vulgaris* var. *Vulgaris*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rudbeckia triloba*, *Schizachyrium scoparium*, *Silphium terebinthinaceum*, *Solidago altissima*, *Solidago gigantea*, *Sorghastrum nutans*, *Symphotrichum laeve*, *Symphotrichum lanceolatum*, *Symphotrichum novae-angliae*, *Symphotrichum pilosum*, *Taraxacum officinale*, *Trifolium hybridum*, *Trifolium pratense*, *Verbascum thapsus*, *Verbena hastata*, *Verbena stricta*, *Viola sororia*, *Zizia aptera*, *Zizia aurea*



Native Species Richness:	11.300
All Species Richness:	14.500
Introduced Species Percent:	21.603
Mean C (Native):	4.014
Mean C (All):	3.229
FQI (Native):	13.566
FQI (All):	11.018
AFQI:	35.870
Shannon Diversity (All):	3.346
Simpson Diversity (All):	0.952
Native Gamma Diversity:	30
All Gamma Diversity:	42

Zone Area (ha):	7.3239
Site Area (with Zone):	24.038
Distance to Chicago (km):	89.0636
Avg. % Rural (8 Buffers):	79.5447
Avg. % Suburban:	7.06662
Avg. % Wilderness:	10.7334
Avg. Ratio of Rural to Suburban (Log):	2.58923
Age of Restoration:	2019
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	Undefined
Ownership Type (PSCC):	Federal

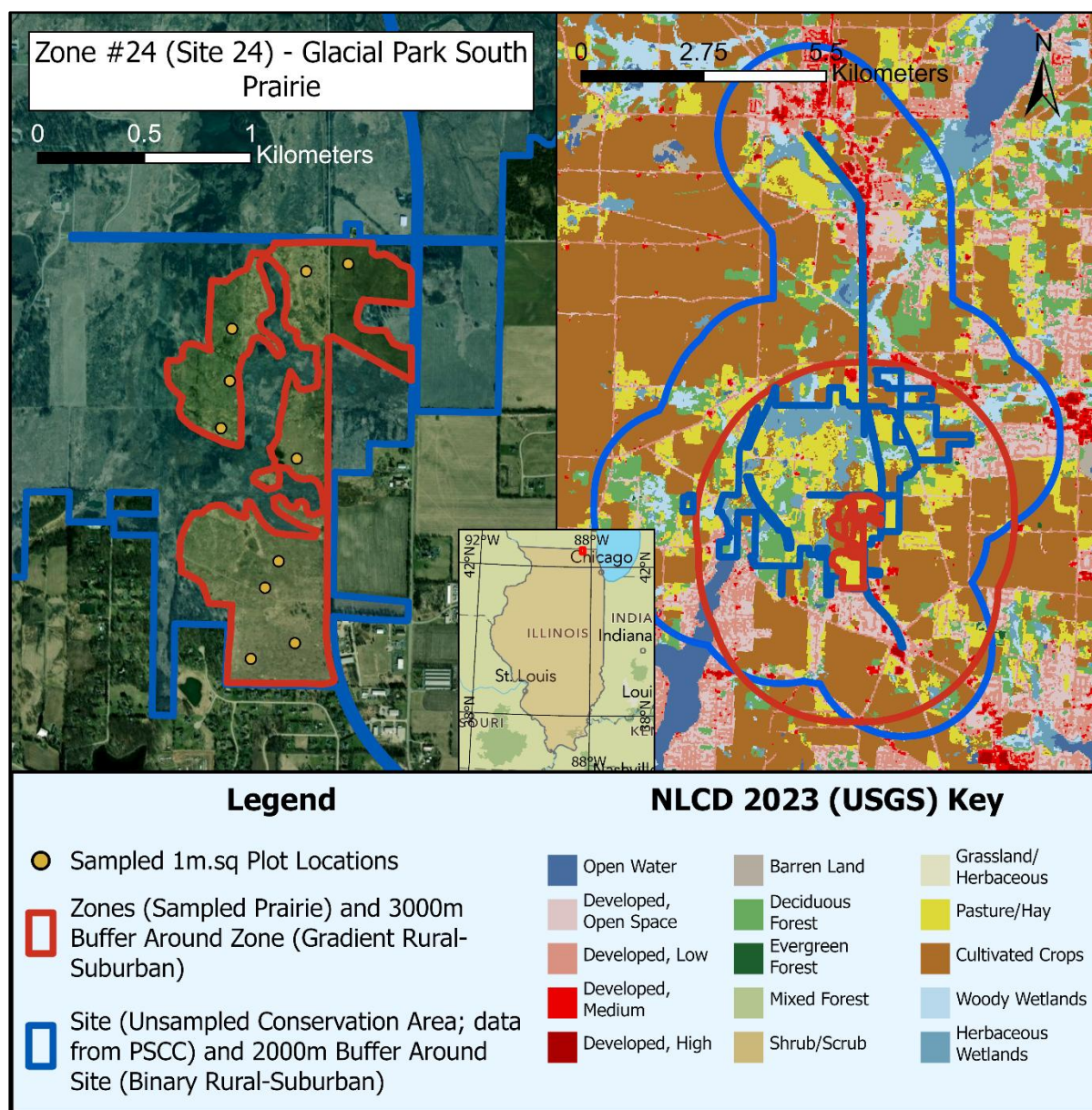
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Zizia aptera*, *Drymocallis arguta*, *Symphyotrichum laeve*, *Heliopsis helianthoides*, *Zizia aurea*,

Top Five Highest Frequency/Coverage Native Species: *Elymus canadensis*, *Zizia aurea*, *Penstemon digitalis*, *Ratibida pinnata*, *Solidago altissima*,

Top Five Highest Frequency/Coverage All Species: *Elymus canadensis*, *Zizia aurea*, *Penstemon digitalis*, *Ratibida pinnata*, *Solidago altissima*,

Zone 24 (Site 24): Glacial Park South Prairie



Approximate Coordinates: 42°24'53"N 88°18'31"W

Date Sampled: September 18-19, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Restoration

Operating Organization: McHenry County Conservation District

Full Species List: *Achillea millefolium*, *Agrostis gigantea*, *Andropogon gerardii*, *Asclepias syriaca*, *Bromus inermis*, *Cirsium altissimum*, *Cornus racemosa*, *Daucus carota*, *Elymus virginicus*, *Eryngium yuccifolium*, *Euthamia graminifolia*, *Geum aleppicum*, *Geum laciniatum*, *Helianthus grosseserratus*, *Melilotus alba*, *Monarda fistulosa*, *Oligoneuron rigidum*, *Pastinaca sativa*, *Penstemon digitalis*, *Phalaris arundinacea*, *Prunella vulgaris* var. *Vulgaris*, *Ratibida pinnata*, *Silphium laciniatum*, *Silphium terebinthinaceum*, *Solidago altissima*, *Sorghastrum nutans*, *Symphyotrichum novae-angliae*, *Symphyotrichum oblongifolium*, *Symphyotrichum pilosum*, *Taraxacum officinale*, *Trifolium hybridum*, *Viola sororia*, *Vitis riparia*



Native Species Richness:	5.600
All Species Richness:	7.900
Introduced Species Percent:	30.547
Mean C (Native):	3.283
Mean C (All):	2.342
FQI (Native):	7.642
FQI (All):	5.528
AFQI:	27.660
Shannon Diversity (All):	2.994
Simpson Diversity (All):	0.928
Native Gamma Diversity:	24
All Gamma Diversity:	33

Zone Area (ha):	104.413
Site Area (with Zone):	1414.67
Distance to Chicago (km):	82.0248
Avg. % Rural (8 Buffers):	33.946
Avg. % Suburban:	19.5603
Avg. % Wilderness:	20.1827
Avg. Ratio of Rural to Suburban (Log):	0.5722
Age of Restoration:	2000
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization Budget Class:	\$5m-\$150m
Ownership Type (PSCC):	Federal

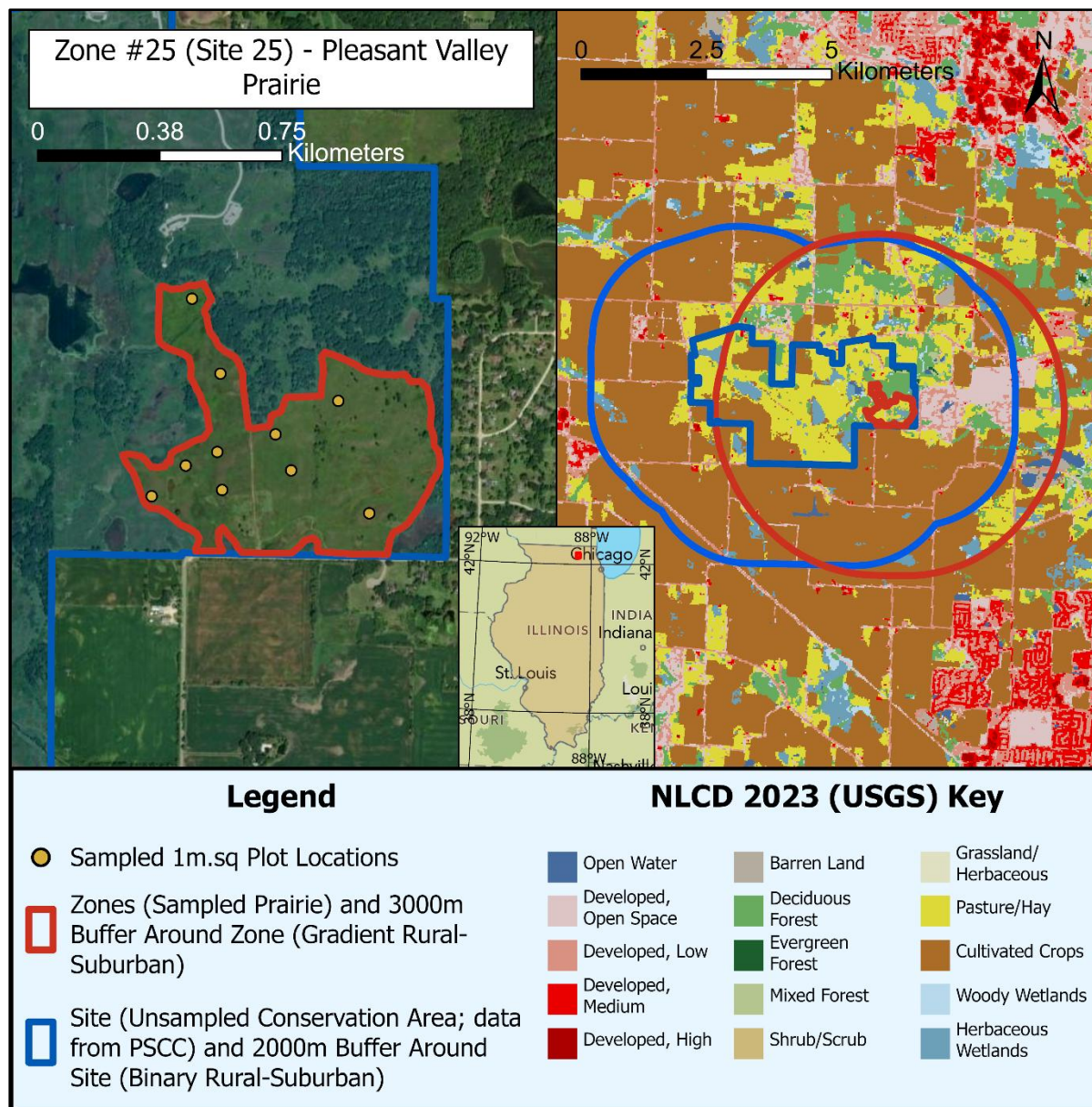
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Symphyotrichum oblongifolium*, *Eryngium yuccifolium*, *Cirsium altissimum*, *Sorghastrum nutans*, *Silphium terebinthinaceum*,

Top Five Highest Frequency/Coverage Native Species: *Sorghastrum nutans*, *Helianthus grosseserratus*, *Solidago altissima*, *Andropogon gerardii*, *Oligoneuron rigidum*,

Top Five Highest Frequency/Coverage All Species: *Bromus inermis*, *Sorghastrum nutans*, *Solidago altissima*, *Helianthus grosseserratus*, *Andropogon gerardii*,

Zone 25 (Site 25): Pleasant Valley Prairie



Approximate Coordinates: 42°14'49"N 88°27'38"W

Date Sampled: September 19, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Mixed

Operating Organization: McHenry County Conservation District

Full Species List: *Achillea millefolium*, *Agrostis gigantea*, *Andropogon gerardii*, *Anemone cylindrica*, *Apocynum cannabinum*, *Bromus inermis*, *Coreopsis palmata*, *Coreopsis tripteris*, *Cornus racemosa*, *Dichanthelium scribnerianum*, *Echinacea pallida*, *Euphorbia corollata*, *Fragaria virginiana*, *Gentiana alba*, *Liatris aspera*, *Melilotus alba*, *Monarda fistulosa*, *Oligoneuron rigidum*, *Pastinaca sativa*, *Prunella vulgaris* var. *Vulgaris*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rhamnus cathartica*, *Rhus glabra*, *Rosa carolina*, *Rudbeckia subtomentosa*, *Schizachyrium scoparium*, *Securigera varia*, *Silphium integrifolium*, *Silphium laciniatum*, *Silphium perfoliatum*, *Silphium terebinthinaceum*, *Solidago altissima*, *Sorghastrum nutans*, *Sporobolus heterolepsis*, *Symphyotrichum lanceolatum*, *Symphyotrichum novae-angliae*, *Symphyotrichum pilosum*, *Viola sororia*, *Vitis riparia*



Native Species Richness:	8.300
All Species Richness:	9.800
Introduced Species Percent:	16.862
Mean C (Native):	3.650
Mean C (All):	3.199
FQI (Native):	10.851
FQI (All):	9.630
AFQI:	34.046
Shannon Diversity (All):	3.190
Simpson Diversity (All):	0.940
Native Gamma Diversity:	33
All Gamma Diversity:	40

Zone Area (ha):	42.8464
Site Area (with Zone):	834.309
Distance to Chicago (km):	77.342
Avg. % Rural (8 Buffers):	27.8729
Avg. % Suburban:	15.8314
Avg. % Wilderness:	28.8553
Avg. Ratio of Rural to Suburban (Log):	0.59249
Age of Restoration:	1984
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization Budget Class:	\$5m-\$150m
Ownership Type (PSCC):	County

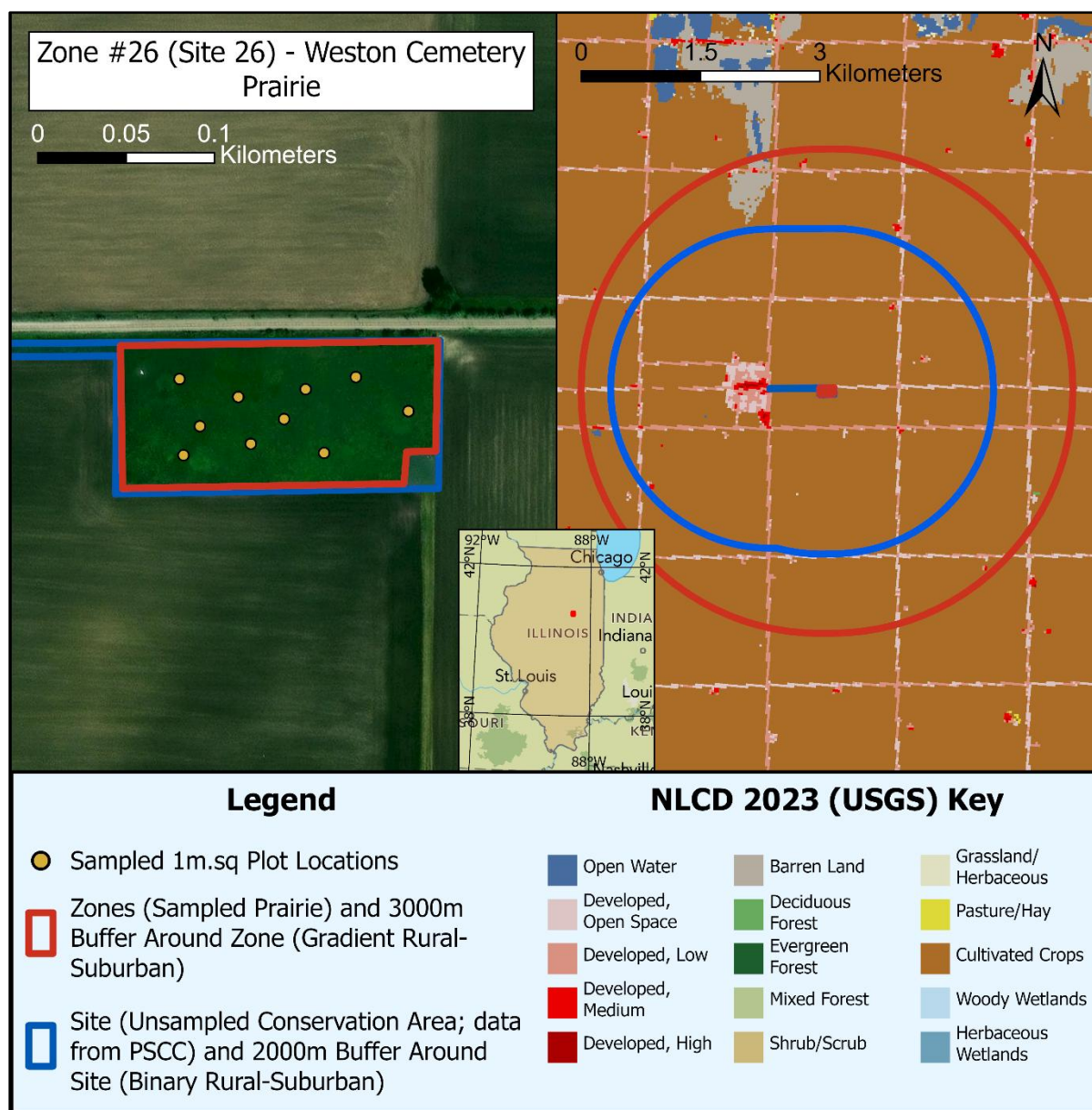
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Echinacea pallida*, *Coreopsis palmata*, *Gentiana alba*, *Rudbeckia subtomentosa*,

Top Five Highest Frequency/Coverage Native Species: *Sorghastrum nutans*, *Solidago altissima*, *Schizachyrium scoparium*, *Ratibida pinnata*, *Silphium terebinthinaceum*,

Top Five Highest Frequency/Coverage All Species: *Sorghastrum nutans*, *Solidago altissima*, *Schizachyrium scoparium*, *Ratibida pinnata*, *Silphium terebinthinaceum*,

Zone 26 (Site 26): Weston Cemetery Prairie



Approximate Coordinates: 40°44'48"N 88°36'49"W

Date Sampled: September 20, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Remnant

Operating Organization: Parklands Foundation

Full Species List: *Agrostis gigantea*, *Amorpha canescens*, *Andropogon gerardii*, *Anemone cylindrica*, *Bromus inermis*, *Ceanothus americanus*, *Cirsium arvense*, *Comandra umbellata*, *Coreopsis palmata*, *Coreopsis tripteris*, *Dichanthelium oligosanthes*, *Echinacea pallida*, *Elymus canadensis*, *Eryngium yuccifolium*, *Euphorbia corollata*, *Euphorbia cyparissias*, *Euthamia graminifolia*, *Fragaria virginiana*, *Gentiana puberulenta*, *Helianthus grosseserratus*, *Helianthus pauciflorus*, *Liatris aspera*, *Lithospermum canescens*, *Lysimachia lanceolata*, *Monarda fistulosa*, *Nabalis asper*, *Oligoneuron rigidum*, *Parthenium integrifolium*, *Pastinaca sativa*, *Phlox pilosa*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rosa carolina*, *Rudbeckia subtomentosa*, *Schizachyrium scoparium*, *Silphium integrifolium*, *Silphium laciniatum*, *Solidago altissima*, *Sorghastrum nutans*, *Sporobolus heterolepsis*, *Symphyotrichum ericoides*, *Symphyotrichum laeve*, *Symphyotrichum oolentangiense*, *Veronicastrum virginicum*, *Viola pedatifida*, *Zizia aurea*



Native Species Richness:	14.700
All Species Richness:	15.800
Introduced Species Percent:	7.188
Mean C (Native):	6.910
Mean C (All):	6.439
FQI (Native):	26.385
FQI (All):	24.679
AFQI:	66.665
Shannon Diversity (All):	3.329
Simpson Diversity (All):	0.953
Native Gamma Diversity:	41
All Gamma Diversity:	46

Zone Area (ha):	1.40522
Site Area (with Zone):	2.02585
Distance to Chicago (km):	143.883
Avg. % Rural (8 Buffers):	89.2228
Avg. % Suburban:	10.6439
Avg. % Wilderness:	0.12452
Avg. Ratio of Rural to Suburban (Log):	2.27061
Age of Restoration:	Remnant
Sector Type (PSCC):	Private
Special Legal Designations:	Yes
Operating Organization	
Budget Class:	< \$5m
Ownership Type (PSCC):	NGO

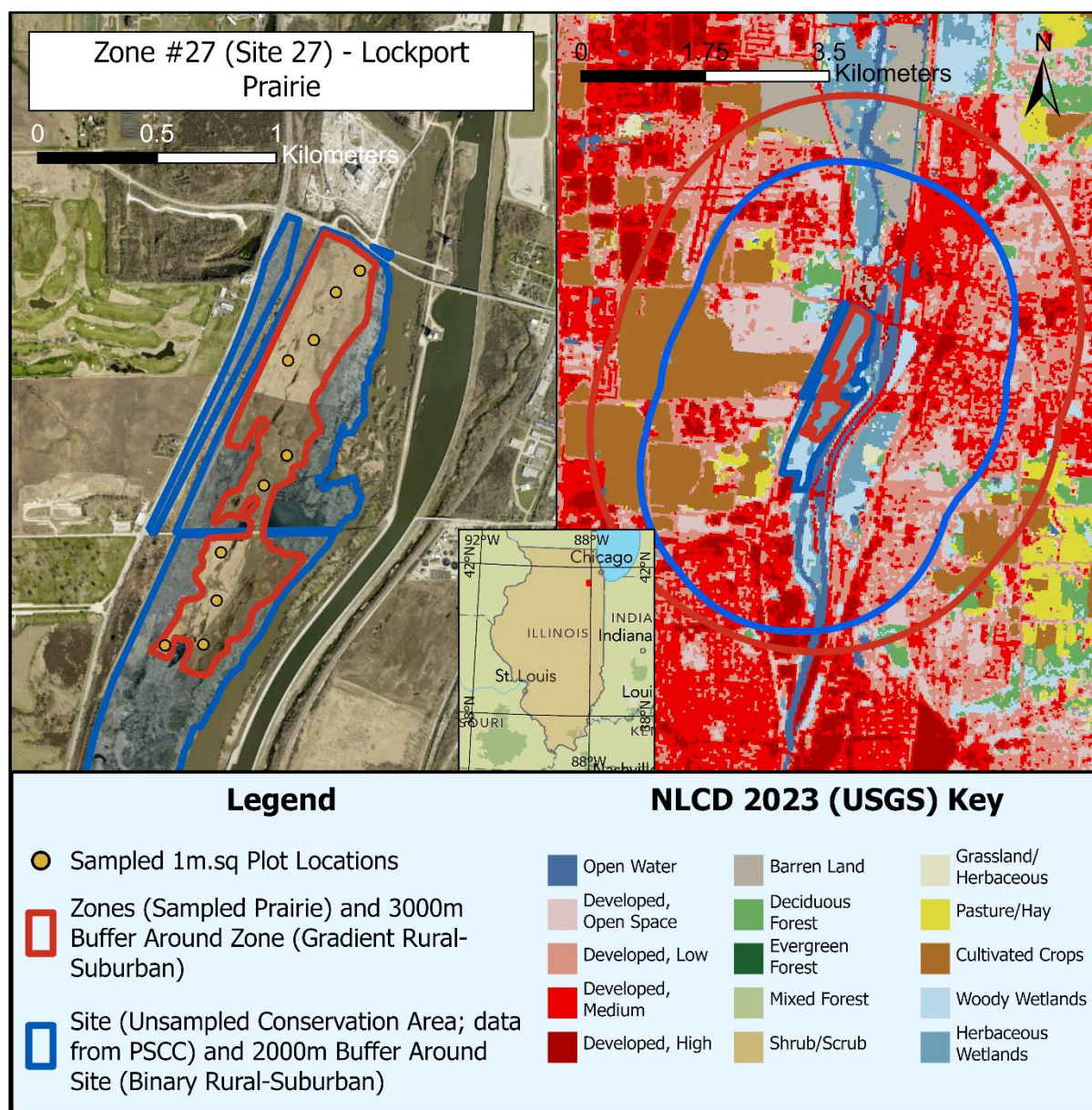
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Nabalus asper*, *Gentiana puberulenta*, *Echinacea pallida*, *Coreopsis palmata*,

Top Five Highest Frequency/Coverage Native Species: *Coreopsis palmata*, *Symphyotrichum laeve*, *Amorpha canescens*, *Symphyotrichum ericoides*, *Sporobolus heterolepsis*,

Top Five Highest Frequency/Coverage All Species: *Coreopsis palmata*, *Bromus inermis*, *Symphyotrichum laeve*, *Amorpha canescens*, *Symphyotrichum ericoides*,

Zone 27 (Site 27): Lockport Prairie



Approximate Coordinates: 41°35'07"N 88°04'31"W

Date Sampled: September 23, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Fully Remnant

Operating Organization: Forest Preserve District of Will County

Full Species List: *Achillea millefolium*, *Ageratina altissima*, *Agrostis gigantea*, *Allium cernuum*, *Ambrosia artemisiifolia*, *Andropogon gerardii*, *Apocynum cannabinum*, *Asclepias verticillata*, *Bouteloua curtipendula*, *Brickellia eupatorioides*, *Bromus inermis*, *Daucus carota*, *Dichanthelium oligosanthes*, *Echinacea purpurea*, *Elymus canadensis*, *Eupatorium altissimum*, *Hypericum sphaerocarpum*, *Lactuca biennis*, *Melilotus alba*, *Monarda fistulosa*, *Nepeta cataria*, *Packera paupercula*, *Pycnanthemum tenuifolium*, *Rhamnus cathartica*, *Schizachyrium scoparium*, *Securigera varia*, *Setaria faberi*, *Solidago altissima*, *Solidago nemoralis*, *Sonchus arvensis*, *Sorghastrum nutans*, *Symphotrichum ericoides*, *Symphotrichum oolentangiense*, *Trifolium hybridum*, *Viola sororia*



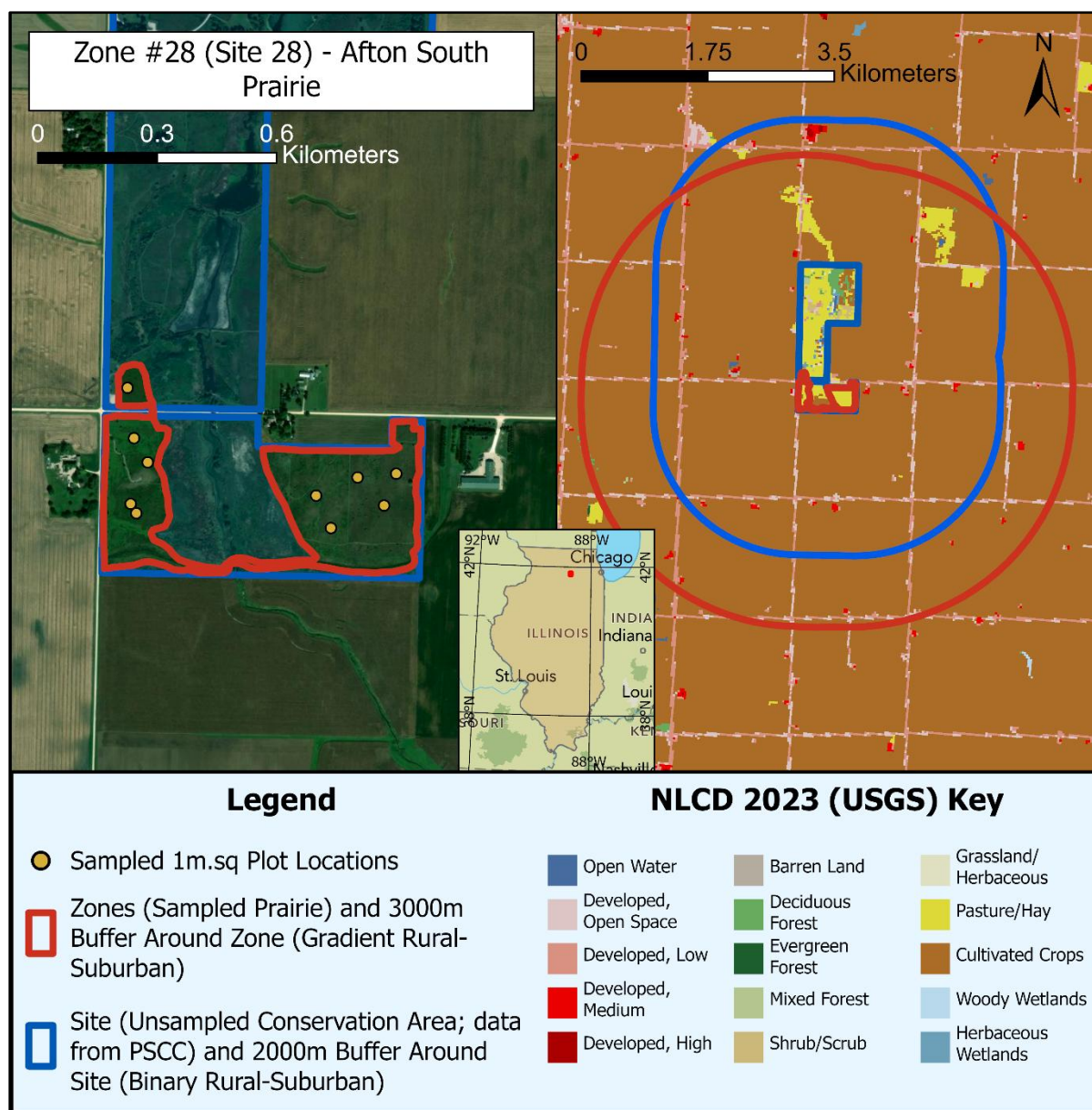
Native Species Richness:	7.300
All Species Richness:	9.600
Introduced Species Percent:	22.865
Mean C (Native):	4.273
Mean C (All):	3.336
FQI (Native):	11.555
FQI (All):	9.169
AFQI:	37.523
Shannon Diversity (All):	3.043
Simpson Diversity (All):	0.927
Native Gamma Diversity:	25
All Gamma Diversity:	35

Zone Area (ha):	48.5592
Site Area (with Zone):	125.629
Distance to Chicago (km):	42.822
Avg. % Rural (8 Buffers):	7.97187
Avg. % Suburban:	45.8399
Avg. % Wilderness:	37.452
Avg. Ratio of Rural to Suburban (Log):	-1.975
Age of Restoration:	Remnant
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	\$5m-
Budget Class:	\$150m
Ownership Type (PSCC):	County

All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Echinacea purpurea*, *Dichanthelium oligosanthes*, *Symphotrichum oolentangiense*, *Bouteloua curtipendula*, *Pycnanthemum tenuifolium*,
Top Five Highest Frequency/Coverage Native Species: *Andropogon gerardii*, *Sorghastrum nutans*, *Monarda fistulosa*, *Solidago altissima*, *Eupatorium altissimum*,
Top Five Highest Frequency/Coverage All Species: *Andropogon gerardii*, *Sorghastrum nutans*, *Monarda fistulosa*, *Solidago altissima*, *Eupatorium altissimum*,

Zone 28 (Site 28): Afton South Prairie



Approximate Coordinates: 41°49'04"N 88°43'57"W

Date Sampled: September 25, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Restoration

Operating Organization: DeKalb County Forest Preserve District

Full Species List: *Achillea millefolium*, *Amorpha canescens*, *Andropogon gerardii*, *Bouteloua curtipendula*, *Bromus inermis*, *Coreopsis tripteris*, *Cornus racemosa*, *Daucus carota*, *Drymocallis arguta*, *Echinacea purpurea*, *Eryngium yuccifolium*, *Gentiana alba*, *Heliopsis helianthoides*, *Hypericum ascyron*, *Lespedeza capitata*, *Lonicera mackii*, *Lupinus polyphyllus*, *Melilotus alba*, *Monarda fistulosa*, *Oligoneuron rigidum*, *Pastinaca sativa*, *Penstemon digitalis*, *Populus tremuloides*, *Prunella vulgaris* var. *Vulgaris*, *Pycnanthemum tenuifolium*, *Ratibida pinnata*, *Schizachyrium scoparium*, *Securigera varia*, *Solidago altissima*, *Solidago speciosa*, *Sonchus arvensis*, *Sorghastrum nutans*, *Sporobolus heterolepis*, *Symphotrichum drummondii*, *Symphotrichum laeve*, *Symphotrichum lanceolatum*, *Symphotrichum novae-angliae*, *Symphotrichum pilosum*, *Trifolium pratense*, *Zizia aurea*



Native Species Richness:	10.100
All Species Richness:	12.000
Introduced Species Percent:	16.471
Mean C (Native):	5.019
Mean C (All):	4.206
FQI (Native):	15.907
FQI (All):	13.441
AFQI:	45.895
Shannon Diversity (All):	3.339
Simpson Diversity (All):	0.953
Native Gamma Diversity:	30
All Gamma Diversity:	40

Zone Area (ha):	17.096
Site Area (with Zone):	124.74
Distance to Chicago (km):	86.6403
Avg. % Rural (8 Buffers):	79.0444
Avg. % Suburban:	8.10526
Avg. % Wilderness:	1.0047
Avg. Ratio of Rural to Suburban (Log):	2.51049
Age of Restoration:	2002
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization	
Budget Class:	< \$5m
Ownership Type (PSCC):	County

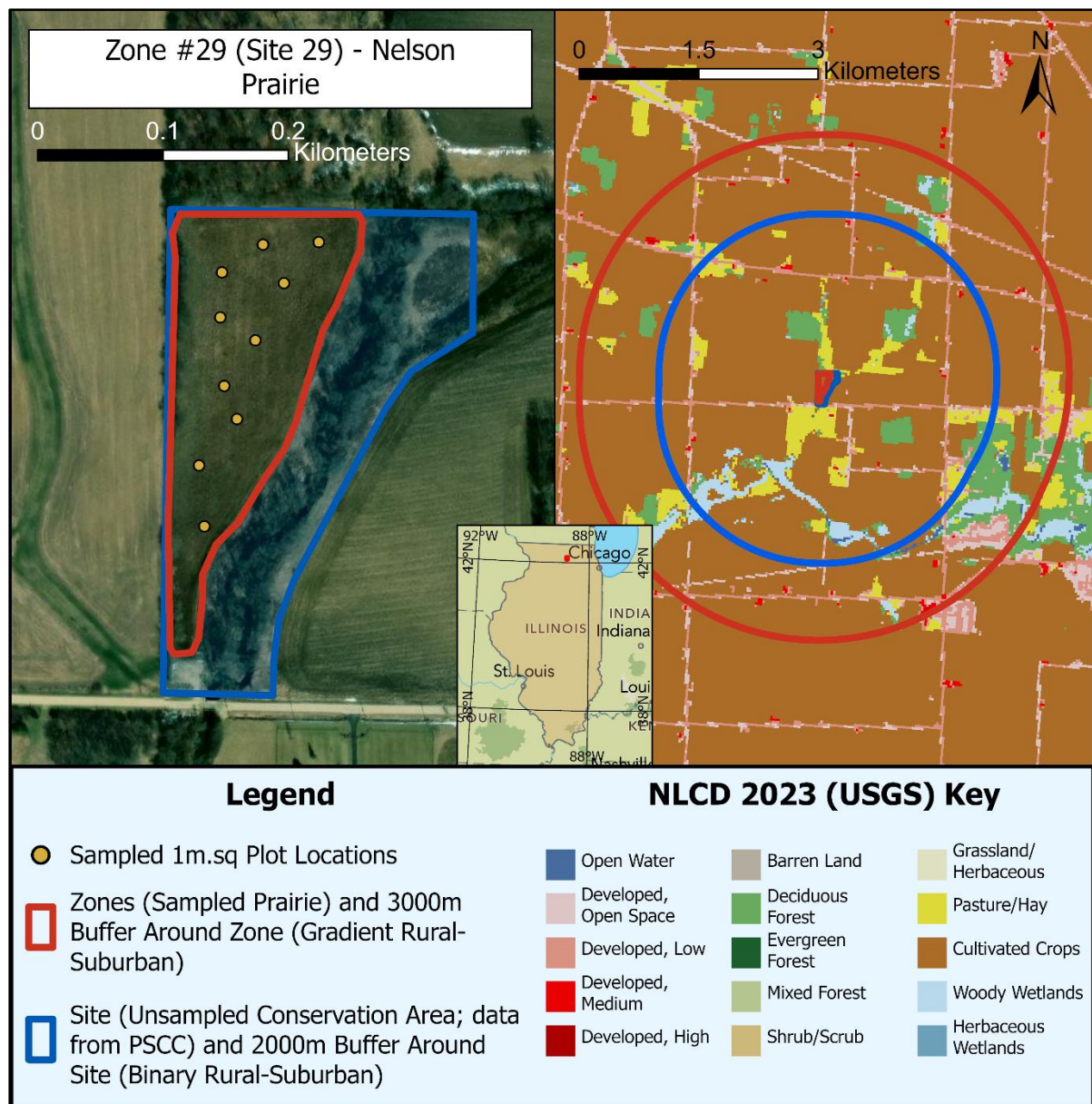
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Sporobolus heterolepsis*, *Hypericum ascyron*, *Echinacea purpurea*, *Drymocallis arguta*, *Amorpha canescens*,

Top Five Highest Frequency/Coverage Native Species: *Sorghastrum nutans*, *Monarda fistulosa*, *Echinacea purpurea*, *Ratibida pinnata*, *Solidago altissima*,

Top Five Highest Frequency/Coverage All Species: *Bromus inermis*, *Sorghastrum nutans*, *Monarda fistulosa*, *Echinacea purpurea*, *Solidago altissima*,

Zone 29 (Site 29): Nelson Prairie



Approximate Coordinates: 42°07'10"N 88°48'07"W

Date Sampled: September 26, 2024

Surrounding Matrix: Rural

Remnant/Restoration Status: Fully Restoration

Operating Organization: DeKalb County Forest Preserve District

Full Species List: *Agrostis gigantea*, *Andropogon gerardii*, *Baptisia alba*, *Bouteloua curtipendula*, *Carex vulpinoidea*, *Chamaecrista fasciculata*, *Cirsium arvense*, *Cornus racemosa*, *Drymocallis arguta*, *Echinacea purpurea*, *Elymus canadensis*, *Eryngium yuccifolium*, *Gaura biennis*, *Gentianella quinquefolia*, *Geum canadense*, *Heliopsis helianthoides*, *Hypericum ascyron*, *Hypericum punctatum*, *Juncus tenuis*, *Lobelia siphilitica*, *Melilotus alba*, *Monarda fistulosa*, *Oenothera biennis*, *Oligoneuron rigidum*, *Penstemon digitalis*, *Prunella vulgaris* var. *Vulgaris*, *Prunus serotina*, *Pycnanthemum tenuifolium*, *Pycnanthemum virginianum*, *Ratibida pinnata*, *Rudbeckia subtomentosa*, *Rumex crispus*, *Schizachyrium scoparium*, *Solidago altissima*, *Solidago gigantea*, *Solidago speciosa*, *Sonchus arvensis*, *Symphyotrichum laeve*, *Symphyotrichum pilosum*, *Taraxacum officinale*, *Trifolium pratense*, *Unidentified Asteraceae* sp., *Verbena stricta*, *Zizia aurea*



Native Species Richness:	19.200
All Species Richness:	23.000
Introduced Species Percent:	16.128
Mean C (Native):	5.024
Mean C (All):	4.217
FQI (Native):	22.022
FQI (All):	18.504
AFQI:	45.934
Shannon Diversity (All):	3.437
Simpson Diversity (All):	0.961
Native Gamma Diversity:	34
All Gamma Diversity:	44

Zone Area (ha):	3.02777
Site Area (with Zone):	6.28956
Distance to Chicago (km):	97.6202
Avg. % Rural (8 Buffers):	74.6993
Avg. % Suburban:	5.52156
Avg. % Wilderness:	7.3857
Avg. Ratio of Rural to Suburban (Log):	2.65107
Age of Restoration:	2019
Sector Type (PSCC):	Public
Special Legal Designations:	No
Operating Organization	
Budget Class:	< \$5m
Ownership Type (PSCC):	County

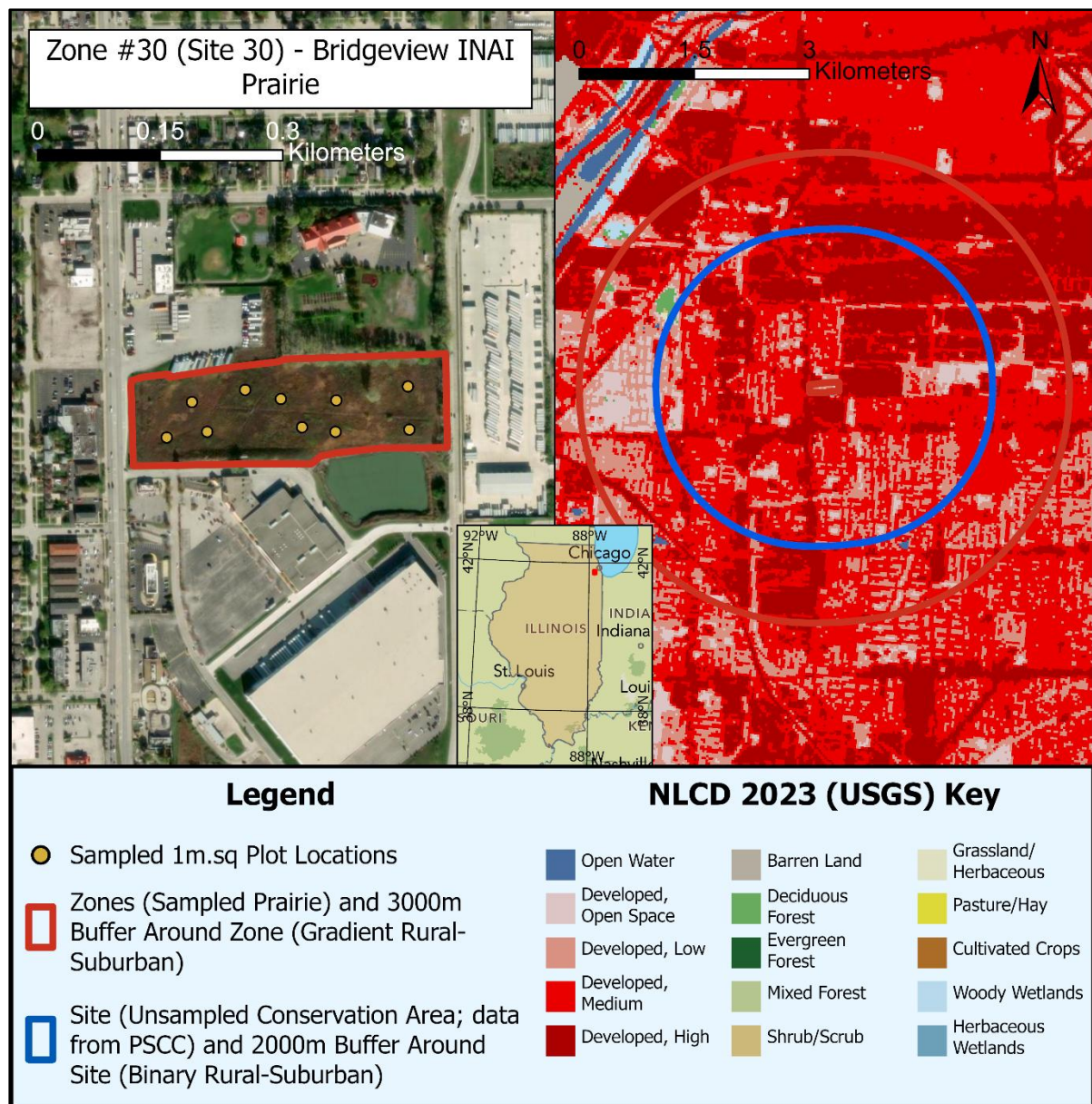
All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Hypericum ascyron*, *Echinacea purpurea*, *Drymocallis arguta*, *Symphyotrichum laeve*, *Eryngium yuccifolium*,

Top Five Highest Frequency/Coverage Native Species: *Penstemon digitalis*, *Solidago altissima*, *Juncus tenuis*, *Symphyotrichum laeve*, *Drymocallis arguta*,

Top Five Highest Frequency/Coverage All Species: *Penstemon digitalis*, *Solidago altissima*, *Juncus tenuis*, *Symphyotrichum laeve*, *Drymocallis arguta*,

Zone 30 (Site 30): Bridgeview INAI Prairie



Approximate Coordinates: 41°45'10"N 87°47'50"W

Date Sampled: September 27, 2024

Surrounding Matrix: Suburban

Remnant/Restoration Status: Mixed

Operating Organization: Village of Bridgeview

Full Species List: *Apocynum cannabinum*, *Asclepias syriaca*, *Baptisia alba*, *Calamagrostis canadensis*, *Eupatorium altissimum*, *Euthamia graminifolia*, *Galium obtusum*, *Geum canadense*, *Helianthus grosseserratus*, *Hypericum perforatum*, *Juncus tenuis*, *Juncus torreyi*, *Lithospermum canescens*, *Lycopus americanus*, *Lycopus uniflorus*, *Lythrum alatum*, *Monarda fistulosa*, *Oenothera biennis*, *Penstemon digitalis*, *Pycnanthemum tenuifolium*, *Pycnanthemum virginianum*, *Rosa carolina*, *Scutellaria parvula*, *Silphium integrifolium*, *Silphium laciniatum*, *Solidago altissima*, *Solidago gigantea*, *Viola sororia*



Native Species Richness:	7.900
All Species Richness:	8.300
Introduced Species Percent:	4.270
Mean C (Native):	4.250
Mean C (All):	4.091
FQI (Native):	11.723
FQI (All):	11.262
AFQI:	41.683
Shannon Diversity (All):	2.927
Simpson Diversity (All):	0.930
Native Gamma Diversity:	27
All Gamma Diversity:	28

Zone Area (ha):	4.4855
Site Area (with Zone):	4.4855
Distance to Chicago (km):	13.15
Avg. % Rural (8 Buffers):	0
Avg. % Suburban:	99.4965
Avg. % Wilderness:	0.10952
Avg. Ratio of Rural to Suburban (Log):	Undefined
Age of Restoration:	1998
Sector Type (PSCC):	Public
Special Legal Designations:	Yes
Operating Organization	\$5m-
Budget Class:	\$150m
Ownership Type (PSCC):	Municipal

All quality and diversity values are the average of ten plot measurements, except for Shannon diversity, Simpson diversity, and both gamma diversities. Landscape and management metrics are assigned at a zone/site level. Years calculated prior to 2024. PSCC stands for the Prairie State Conservation Coalition.

Top Five C-Value Species: *Scutellaria parvula*, *Lithospermum canescens*, *Baptisia alba*, *Pycnanthemum tenuifolium*, *Lythrum alatum*,

Top Five Highest Frequency/Coverage Native Species: *Calamagrostis canadensis*, *Euthamia graminifolia*, *Solidago gigantea*, *Juncus tenuis*, *Solidago altissima*,

Top Five Highest Frequency/Coverage All Species: *Calamagrostis canadensis*, *Euthamia graminifolia*, *Solidago gigantea*, *Juncus tenuis*, *Solidago altissima*,

Appendix F: Full Coding Scripts

Python-Based Land Cover Analysis

This includes the functions used to perform data exploration and extraction from the National Land Cover Database in order to measure percentages and areas of the landscape around a given zone or site. It is working from shapefiles that have already partially been cleaned, projected, and prepared by hand in ArcGIS Pro. It does not include the code itself used to run and implement these functions, since that is fairly straightforward to do and largely redundant.

```
Nlcd_matrix_analysis_functions.py
# Andy Sima - Nov-Jan 2024-2025 (Copied June 2 2025)
# Master's Thesis Analysis Work
# Analyzing the National Land Cover Database for Landscape Matrix cover
# https://www.usgs.gov/centers/eros/science/annual-national-land-cover-
database
# https://www.usgs.gov/centers/eros/science/annual-nlcd-science-product-
user-guide

# DEFINING FUNCTIONS FOR LANDSCAPE MATRIX ANALYSIS
# Version 1.1 - Dated 01/22/2025

# ## LIBRARIES NECESSARY FOR THESE FUNCTIONS ## #
import datetime
from typing import Dict, List
import pandas as pd
import os
from pathlib import Path
import arcpy
arcpy.CheckOutExtension("Spatial")
arcpy.CheckOutExtension("Image")

# ## ALL THE INFORMATION IN ONE CONVENIENT PLACE ## #
# ## TABLE OF CONTENTS ## #

'''
0. Initialization Variables as utilized in running the script.
1. Fixing Filepaths slash_fix_to_dict():
2. Getting runtime strings time_from_start():
3. Fixing Filepath end slashes filepath_end_slash():
4. Getting Site Number get_site_number():
5. Getting the Buffer size get_buff_size():
6. Deleting from a list delete_from_list():
7. Selecting sites and buffering them individually
select_and_buffer_many():
8. Extracting the land cover data land_cover_data_extract():
9. Creating the index name based on filepath get_index_name():
10. Extracting rows from ArcGIS to a pandas Dataframe
extract_multi_row_to_df():
11. Finding what percent each class has find_percents_of_area():
12. Extracting Confidence Intervals from arcgis
extract_confidence_interval():
13. Streamlining the implementation of all these functions
streamlined_nlcd_analysis():
14. Changelog, Backups, and old versions of code.
'''
```

```

'''
# ## 0. Initialization Variables ## #
'''

# This section represents an EXAMPLE OF the crucial global variables needed
# to make
# this script run properly. That means the things like the filepaths, the
# classes
# keys, the buffer distances, etc. Make sure you update the filepaths
# before using
# these initialization variables.
# It also includes an example of the streamlined function usage, with all
# arguments.

# ## #
# from typing import Dict, List
# from nlcd_matrix_analysis_functions import *
# import os
# import pandas as pd
# from pathlib import Path
#
# import arcpy
# arcpy.CheckOutExtension("Spatial")
# arcpy.CheckOutExtension("Image")
# print(f"Successfully loaded libraries in
# {time_from_start(processStartTime)}\n")
# print("Now running analysis...\n")
#
#
# # # DEFINING COORDINATE SYSTEM
# WKT_coord_sys =
# 'PROJCS["NAD_1983_2011_UTM_Zone_16N",GEOGCS["GCS_NAD_1983_2011",DATUM["D_NA
# D_1983_2011",SPHEROID["GRS_1980",6378137.0,298.257222101]],PRIMEM["Greenwic
# h",0.0],UNIT["Degree",0.0174532925199433]],PROJECTION["Transverse_Mercator"
# ],PARAMETER["False_Easting",500000.0],PARAMETER["False_Northing",0.0],PARAM
# ETER["Central_Meridian",-
# 87.0],PARAMETER["Scale_Factor",0.9996],PARAMETER["Latitude_Of_Origin",0.0],
# UNIT["Meter",1.0]]'
# coord_sys = arcpy.SpatialReference(6345)
#
# # # # DEFINING FILEPATHS ## #
# filepaths_list_first = [("workspace_filepath",
# r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
# (ansi0618)\gis_BIGDATA_files\\"),
# #
# # ("scratch_filepath",
# r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
# (ansi0618)\gis_BIGDATA_files\scratch\\"),
# #
# # ("local_scratch",
# r"C:\Users\ansi0618\this_pc_storage\local_gis_scratch"),
# #
# #
# # ("sites_polygon_filepath",
# r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS\Stockholm Classes\THESIS
# WORK\GIS_resources\vector_data\all_areas_PSCC23_utml6n_dslve_final"),
# #
# # ("zones_polygon_filepath",
# r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS\Stockholm Classes\THESIS
# WORK\GIS_resources\vector_data\zones_of_interest.shp"),
# #
# #
# # ("nlcd_processed_filepath_10",
# r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
# (ansi0618)\gis_BIGDATA_files\NLCD_Illinois_2023\processed_files\Annual_NLCD
# _LndCov_2023_CU_C1V0_UTM16N_clip50mi_mask_LEVEL0.tif"),

```

```

#                                     ("nlcd_processed_filepath_l1",
r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
(ansi0618)\gis_BIGDATA_files\NLCD_Illinois_2023\processed_files\Annual_NLCD
_LndCov_2023_CU_C1V0_UTM16N_clip50mi_mask_LEVEL1.tif"),
#                                     ("nlcd_processed_filepath_l2",
r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
(ansi0618)\gis_BIGDATA_files\NLCD_Illinois_2023\processed_files\Annual_NLCD
_LndCov_2023_CU_C1V0_UTM16N_clip50mi_mask_LEVEL2.tif"),
#                                     ("nlcd_processed_filepath_l3",
r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
(ansi0618)\gis_BIGDATA_files\NLCD_Illinois_2023\processed_files\Annual_NLCD
_LndCov_2023_CU_C1V0_UTM16N_clip50mi_mask_LEVEL3.tif"),
#
#                                     ("land_cover_polygons_L0",
r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
(ansi0618)\gis_BIGDATA_files\NLCD_Illinois_2023\processed_files\NLCD_landCo
ver_2023_polygons_clip5km_LEVEL0.shp"),
#                                     ("land_cover_polygons_L1",
r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
(ansi0618)\gis_BIGDATA_files\NLCD_Illinois_2023\processed_files\NLCD_landCo
ver_2023_polygons_clip5km_LEVEL1.shp"),
#                                     ("land_cover_polygons_L2",
r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
(ansi0618)\gis_BIGDATA_files\NLCD_Illinois_2023\processed_files\NLCD_landCo
ver_2023_polygons_clip5km_LEVEL2.shp"),
#                                     ("land_cover_polygons_L3",
r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
(ansi0618)\gis_BIGDATA_files\NLCD_Illinois_2023\processed_files\NLCD_landCo
ver_2023_polygons_clip5km_LEVEL3.shp"),
#
#                                     ("confidence_filepath",
r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
(ansi0618)\gis_BIGDATA_files\NLCD_Illinois_2023\processed_files\Annual_NLCD
_LndCnf_2023_CU_C1V0_UTM16N_clip50mi_mask_5mibuff.tif"),
#
#                                     ("table_output_filepath",
r"U:\Kursbiblioteket\GE9011\student_files\Sima Andrew David
(ansi0618)\gis_BIGDATA_files\outputs\\")
#
# filepaths_list = slash_fix_to_dict(filepaths_list_first)
#
# # ## DEFINING ARCPY ENVIRONMENTAL SETTINGS ## #
# arcpy.env.workspace = filepaths_list["workspace_filepath"]
# arcpy.env.scratchWorkspace = filepaths_list["scratch_filepath"]
# arcpy.env.outputCoordinateSystem = coord_sys
# print(f"Output Coordinate System is: {coord_sys.name}\n")
#
# ## DEFINING GLOBAL VARIABLES ## #
# sites_list = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18,
19,
#                 22, 23, 24, 25, 26, 27, 28, 29, 30]
# zones_list = list(range(1, 31))
# buff_dist_list_sites = [2000]
# buff_dist_list_zones = [50, 200, 500, 1000, 1500, 2000, 2500, 3000]
#
# land_cover_dict = {11: {"L0": "Open Water", "L1": [1, "Water"], "L2": [1,
"Water"],
#
#                                     "L3": [1, "Water"]},
#
#                 12: {"L0": "Perennial Ice/Snow", "L1": [1, "Water"],
"L2": [1, "Water"],
#
#                                     "L3": [1, "Water"]},

```



```

#           21: {"L0": "Developed, Open Space", "L1": [2,
"OpenLandDev"], "L2": [2, "Developed"],
#           "L3": [2, "Developed"]},
#           22: {"L0": "Developed, Low Intensity", "L1": [3,
"LowDev"], "L2": [2, "Developed"],
#           "L3": [2, "Developed"]},
#           23: {"L0": "Developed, Medium Intensity", "L1": [3,
"LowDev"], "L2": [2, "Developed"],
#           "L3": [2, "Developed"]},
#           24: {"L0": "Developed, High Intensity", "L1": [4,
"HighDev"], "L2": [2, "Developed"],
#           "L3": [2, "Developed"]},
#           31: {"L0": "Barren Land (Rock/Soil/Clay)", "L1": [5,
"BarrenLand"], "L2": [3, "Ecosystem"],
#           "L3": [4, "OpenEcosystem"]},
#           41: {"L0": "Deciduous Forest", "L1": [6,
"ForestEcosystem"], "L2": [3, "Ecosystem"],
#           "L3": [3, "ClosedEcosystem"]},
#           42: {"L0": "Evergreen Forest", "L1": [6,
"ForestEcosystem"], "L2": [3, "Ecosystem"],
#           "L3": [3, "ClosedEcosystem"]},
#           43: {"L0": "Mixed Forest", "L1": [6,
"ForestEcosystem"], "L2": [3, "Ecosystem"],
#           "L3": [3, "ClosedEcosystem"]},
#           52: {"L0": "Shrub/Scrub", "L1": [7, "GrassEcosystem"],
"L2": [3, "Ecosystem"],
#           "L3": [3, "ClosedEcosystem"]},
#           71: {"L0": "Grassland/Herbaceous", "L1": [7,
"GrassEcosystem"], "L2": [3, "Ecosystem"],
#           "L3": [4, "OpenEcosystem"]},
#           81: {"L0": "Pasture/Hay", "L1": [8, "Field"], "L2":
[5, "OpenField"],
#           "L3": [6, "OpenField"]},
#           82: {"L0": "Cultivated Crops", "L1": [9, "Farmland"],
"L2": [4, "Farmland"],
#           "L3": [5, "Farmland"]},
#           90: {"L0": "Woody Wetlands", "L1": [10,
"WetEcosystem"], "L2": [3, "Ecosystem"],
#           "L3": [3, "ClosedEcosystem"]},
#           95: {"L0": "Emergent Herbaceous Wetlands", "L1": [10,
"WetEcosystem"], "L2": [3, "Ecosystem"],
#           "L3": [4, "OpenEcosystem"]}
#           }
#
# reclassify_classes_list = {"L0": {11: "Open Water", 12: "Perennial
Ice/Snow", 21: "Developed, Open Space",
#           22: "Developed, Low Intensity", 23:
"Developed, Medium Intensity",
#           24: "Developed, High Intensity", 31:
"Barren Land (Rock/Soil/Clay)",
#           41: "Deciduous Forest", 42: "Evergreen
Forest", 43: "Mixed Forest",
#           52: "Shrub/Scrub", 71:
"Grassland/Herbaceous", 81: "Pasture/Hay",
#           82: "Cultivated Crops", 90: "Woody
Wetlands", 95: "Emergent Herbaceous Wetlands"},
#           "L1": {1: "Water", 2: "OpenLandDev", 3:
"LowDev", 4: "HighDev",
#           5: "BarrenLand", 6: "ForestEcosystem",
7: "GrassEcosystem",

```

```

#                                     8: "Field", 9: "Farmland", 10:
"WetEcosystem"},
#                                     "L2": {1: "Water", 2: "Developed", 3:
"Ecosystem", 4: "Farmland", 5: "OpenField"},
#                                     "L3": {1: "Water", 2: "Developed", 3:
"ClosedEcosystem", 4: "OpenEcosystem",
#                                     5: "Farmland", 6: "OpenField"}}
#
# classes_key_list = {"L0": [11, 12, 21, 22, 23, 24, 31, 41, 42, 43, 52,
71, 81, 82, 90, 95],
#                     "L1": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10],
#                     "L2": [1, 2, 3, 4, 5],
#                     "L3": [1, 2, 3, 4, 5, 6]
#                     }
#
# # An example of the streamlined function in use with all arguments, for
# # "Sites" analysis at a "Level 0" classification scheme.
#
#
streamlined_nlcd_analysis(in_sites_file=filepaths_list["sites_polygon_filep
ath"],
#                           buffer_dist_list=buff_dist_list_sites,
#
in_nlcd_polygon=filepaths_list["land_cover_polygons_L0"],
#
scratch_filepath=filepaths_list["scratch_filepath"],
#
output_filepath=filepaths_list["table_output_filepath"],
#                           site_iterate_list=sites_list,
#                           landuse_col_keys=classes_key_list["L0"],
#                           coord_sys=coord_sys,
#
confidence_filepath=filepaths_list["confidence_filepath"],
#                           new_filename="sites_level0_full_binary",
#                           erase_siteyn=True,
#                           delete_yn=True,
#                           new_col=True)
# ## #

```

```

'''
## 1. FIXING FILEPATHS ## #
'''
# This function fixes the weird backslashes that copied Windows filepaths
all have
# and replaces them with more ArcPro-friendly forwardslash versions of
themselves.
# It requires a list of all the variables/filepaths you want fixed, and
returns a
# dictionary where each fixed filepath is keyed to a string version of its
# original variable name.
# The input list must be formatted as a list of lists with "name" and
"filepath" (value)
# pairs for this function to work.

```

```

def slash_fix_to_dict(in_list):
    fixed_dict = {}

    for name, value in in_list:
        fixed_value = value.replace("\\", "/").replace("//", "/")

```

```

        fixed_dict[name] = fixed_value

    return fixed_dict

'''
# ## 2. GETTING RUNTIME STRINGS ## #
'''
# This function uses datetime.datetime.now() objects to present the time
# that's
# passed between two points in time, as a string format. It requires a
# datetime
# variable as the starting point, and defaults to the moment it's called
# for the
# endpoint. It returns a time as a string.

def time_from_start(start, now=None):
    if now is None:
        time_from = datetime.datetime.now() - start
    else:
        time_from = now - start
    return str(time_from)

'''
# ## 3. FIXING FILEPATH END SLASHES ## #
'''
# ## OR: VALIDATE THE SCRATCH WORKSPACE FORMATTING ## #
# This function makes sure that the scratch workspace filepath, and other
# applicable filepaths, end in a "/" to distinguish them from actual files.
# Requires a string and returns a string.

def filepath_end_slash(in_path):
    if in_path[-1] != "/":
        return in_path + "/"
    else:
        return in_path

'''
# ## 4. GETTING THE SITE NUMBER ## #
'''
# This function finds the part of the filename that includes the site or
# zoi number
# and returns it as a string. It excludes for generic "sites" in the
# filepath,
# but you have to be careful about how your filename is structured going
# in.
# It defaults to look for both site or zoi, with _ as the delimiter, but
# you can
# change that accordingly.
# It requires the in_text file string and returns one part of the file name
# with the
# site number.

def get_site_number(in_text, tag1="site", tag2="zoi", exclude="sites",
delim="_"):

```



```

parts = in_text.split(delim)
for x in parts:
    if tag1 in x and exclude not in x:
        return x
    elif tag2 in x and exclude not in x:
        return x

    raise ValueError(f"Text tag \"{tag1}\" or \"{tag2}\" not found in
filename.")

'''
# ## 5. GETTING THE BUFFER SIZE ## #
'''
# This function finds the part of the filename that includes the buffer
distance.
# It works almost identically to get_site_number, and honestly I don't know
why
# I made two of these. Oh well. It just only checks for one tag instead of
two.
# It requires the in_text file string and returns the part of the file
string with
# the buff size as a string.

def get_buff_size(in_text, tag="buff", delim="_"):
    parts = in_text.split(delim)
    for x in parts:
        if tag in x:
            return x

    raise ValueError(f"Text tag \"{tag}\" not found in filename.")

'''
# ## 6. DELETING ALL ARCGIS THINGS FROM A LIST ## #
'''
# This function iteratively deletes all files related to ArcGIS
infrastructure
# when it is passed a list of filepaths. It does not return anything and
# only requires a list of filepaths.

def delete_from_list(files=None):
    if files is None:
        files = []

    print(f"Deleting {str(len(files))} scratch files now...")

    for filei in files:
        arcpy.management.Delete(
            in_data=filei,
            data_type=""
        )

'''
# ## 7. SELECTING SITES AND BUFFERING THEM INDIVIDUALLY ## #
'''
# This function takes the original file with all combined sites/zois and
separates

```

```

# them into singular sites, and then buffers them, as many times as
distances are given
# in the function, with the option to then erase the original shape of the
site,
# leaving only the buffer behind. This is preferred for analysis as it only
focuses on
# land cover beyond the scope of the site.
# It also returns a list of all the clipped/erased files created.
# This function requires a file of sites, the workspace to send them to,
and the list
# of site numbers/keys to iterate over.
# in_file is the sites/zois to be buffered, out_folder is the scratch space
to
# send them (must end in an /), iterate_list is the list of site
numbers/keys that the
# function will use to iterate its site distinctions, arc_iterate_col is
the name of the
# column in ArcMap that contains the iterate keys for the file, buff_dist
is a
# list of the distances you want to buffer the sites, site_or_zoi is
whether you want
# the tag to be "site" or something else, erase_site_yn is whether or not
you want
# to erase the original site footprint from the buffer, and delete_step_yn
is whether
# you want to delete all the step files created in the process.

```

```

def select_and_buffer_many(in_file, out_folder, iterate_list,
arc_iterate_col="site_num",
buff_dist=None, site_or_zoi="site",
erase_site_yn=True,
delete_step_yn=False):

    buff_list = []
    if buff_dist is None:
        buff_dist = [2000]
    for j in buff_dist:
        buff_list.append(j)

    to_list_append = []
    del_list = []

    stripped_infile = os.path.splitext(os.path.basename(in_file))[0]
    out_folder_fixed = filepath_end_slash(out_folder)

    for it in iterate_list:

        selected_filepath =
f"{out_folder_fixed}{stripped_infile}_{site_or_zoi}{it}"
        selected_filepathshp = selected_filepath + ".shp"

        arcpy.conversion.ExportFeatures(
            in_features=in_file,
            out_features=selected_filepathshp,
            where_clause=f"{arc_iterate_col} = " + str(it),
            use_field_alias_as_name="NOT_USE_ALIAS",
        )

    for h in buff_list:

```

```

buffer_filepath = selected_filepath + f"_buff{h}m"
buffer_filepathshp = buffer_filepath + ".shp"

arcpy.analysis.PairwiseBuffer(
    in_features=selected_filepathshp,
    out_feature_class=buffer_filepathshp,
    buffer_distance_or_field=f"{str(h)} Meters",
    dissolve_option="ALL",
    dissolve_field=None,
    method="PLANAR",
    max_deviation="0 Meters"
)

if erase_site_yn is True:
    clip_filepath = buffer_filepath + "_erased"
    clip_filepathshp = clip_filepath + ".shp"

    arcpy.analysis.PairwiseErase(
        in_features=buffer_filepathshp,
        erase_features=selected_filepathshp,
        out_feature_class=clip_filepathshp,
        cluster_tolerance=None
    )
    print(f"Saved buffered and erased file to
{clip_filepathshp}")

    del_list.append(buffer_filepathshp)
    del_list.append(selected_filepathshp)
    to_list_append.append(clip_filepathshp)

else:
    print(f"Saved buffered file to {buffer_filepathshp}")

    del_list.append(selected_filepathshp)
    to_list_append.append(buffer_filepathshp)

if delete_step_yn is True:
    delete_from_list(del_list)

return to_list_append

'''
# ## 8. EXTRACTING THE LAND COVER DATA ## #
'''
# This function takes a buffered shapefile, clips the larger NLCD polygons
to just
# the extent of the buffer, and then calculates the area geometry of each
type of NLCD
# polygon within just the clipped buffer.
# It requires a list of shapefiles and the NLCD polygon buffer, creates new
individual
# polygon shapefiles, and returns a list of the names of the new polygon
shapefiles.
# In_buffer is the [list] of buffers to extract data from, nlcd_polyfile is
the
# master nlcd polygon, out_folder is the scratch space to save the
nlcd_shapefiles,
# coordinate_system is the coordinate system, area_field_name is what you
want to name

```



```
# the new field during calculate geometry, area_unit is what unit you want
measurements
# to be in. There is no delete option since this tool produces no step
files.
```

```
def land_cover_data_extract(in_buffer, nlcd_polyfile, out_folder,
                             coordinate_system,
                             area_field_name="area_ha",
                             area_unit="HECTARES"):
    to_list_append = []

    stripped_nlcd = os.path.splitext(os.path.basename(nlcd_polyfile))[0]
    level_tag = get_buff_size(stripped_nlcd, tag="LEVEL")

    for buffer in in_buffer:

        stripped_inbuffer = os.path.splitext(os.path.basename(buffer))[0]
        out_file = out_folder + f"{stripped_inbuffer}_NLCDclip_{level_tag}"
        out_polyshp = out_file + ".shp"
        sitenum = get_site_number(stripped_inbuffer)
        buff_size = get_buff_size(stripped_inbuffer)

        print(f"Preparing to extract {level_tag} land cover data for
        {sitenum}_{buff_size}...")

        arcpy.analysis.PairwiseClip(
            in_features=nlcd_polyfile,
            clip_features=buffer,
            out_feature_class=out_polyshp,
            cluster_tolerance=None
        )

        arcpy.management.AddField(
            in_table=out_polyshp,
            field_name=area_field_name,
            field_type="FLOAT",
            field_precision=None,
            field_scale=None,
            field_length=None,
            field_alias="",
            field_is_nullable="NULLABLE",
            field_is_required="NON_REQUIRED",
            field_domain=""
        )

        arcpy.management.CalculateGeometryAttributes(
            in_features=out_polyshp,
            geometry_property=f"{area_field_name} AREA",
            length_unit="",
            area_unit=area_unit,
            coordinate_system=coordinate_system,
            coordinate_format="SAME_AS_INPUT"
        )

        to_list_append.append(out_polyshp)

        print(f"File saved to: {out_polyshp}")

    print(f"Finished clipping and extracting for {str(len(to_list_append))}
    sites.\n")
```

```

    return to_list_append

'''
# ## 9. CREATING THE INDEX NAME BASED ON FILEPATH ## #
'''
# This function uses previously established functions to generated a unique
# index_name variable based off of the input files existing filepath tags.
# It requires a valid string filepath that conforms to existing index
naming
# conventions, and returns a string as an index name representative of the
file
# being worked on in this context.

def get_index_name(in_file):
    stripped_infile = os.path.splitext(os.path.basename(in_file))[0]
    sitenum = get_site_number(stripped_infile)
    buffsize = get_buff_size(stripped_infile)
    leveltag = get_buff_size(stripped_infile, tag="LEVEL")
    index_name = f"{sitenum}_{buffsize}_{leveltag}"

    return index_name

'''
# ## 10. EXTRACTING ROWS FROM ARCGIS TO A PANDAS DATAFRAME ## #
'''
# This function takes a list of input arcgis files and extracts the
relevant
# attributes from the attribute table to a pandas dataframe. It does not
perform
# any sort of calculations on the dataframe or the attributes themselves.
# it requires a list of arcgis files to analyse, and it requires a list of
keys
# that you want it to recognize as unique columns (i.e. land cover codes).
It also
# optionally requests a list of columns to extract from within arcgis, if
you want
# to extract more than just "area_ha."
# It returns a pandas dataframe with extract values from every input file.

def extract_multi_row_to_df(filepath, col_keys, data_extract_field=None,
                             base_fields="gridcode"):

    if data_extract_field is None:
        data_extract_field = ["area_ha"]

    return_df = pd.DataFrame(columns=col_keys)

    for file in filepath:
        index_name = get_index_name(file)

        print(f"Extracting Rows from {index_name}...")

        outer_df = pd.DataFrame(columns=col_keys)

        for extract_field in data_extract_field:
            fields_list = [base_fields, extract_field]
            extract_field_dict = {}

```

```

        with arcpy.da.SearchCursor(file, fields_list) as cursor:
            for row in cursor:
                extract_field_dict[row[0]] = row[1]

                site_row_name = f"{index_name}_{extract_field}"

                extract_df = pd.DataFrame(extract_field_dict,
index=[site_row_name])
                outer_df = pd.concat([outer_df, extract_df])

                return_df = pd.concat([return_df, outer_df])

return_df_0 = return_df.fillna(0)

return return_df_0

'''
# ## 11. FINDING WHAT PERCENTAGES EACH CLASS HAS ## #
'''
# This function takes a pandas dataframe and summarizes each row so it
produces
# a second row displaying what percentage each term in the first row
represents
# of the entire whole. I.e. if you have a row of areas as input, this
function will
# concatenate new rows onto your dataframe that tell you what percent of
the total
# area each individual term represents. It also has an optional function to
produce a
# new column with the total area of each row.
# This function requires a pandas dataframe whose cell values are in a
number format.
# it returns the input dataframe, now with new rows of % values.

def find_percents_of_area(in_df, new_row_tag="%area", new_column_yn=False,
                        new_col_tag="total_area"):
    row_list = in_df.index.tolist()

    if new_column_yn is True:
        in_df[new_col_tag] = 0

    for row in row_list:
        total_area = in_df.loc[row].sum()
        percent_dict = {}

        for column in in_df.columns:
            percent_dict[column] = (in_df[column].loc[row] / total_area) *
100

        if new_column_yn is True:
            in_df.loc[row, new_col_tag] = total_area
            percent_dict[new_col_tag] = total_area

        percent_df = pd.DataFrame(percent_dict,
index=[f"{row}_{new_row_tag}"])
        in_df = pd.concat([in_df, percent_df])

    in_df_final = in_df.sort_index()

```



```

    return in_df_final

'''
# ## 12. EXTRACTING CONFIDENCE INTERVALS FROM ARCGIS ## #
'''
# This function produces a table that takes the average and standard
deviation of
# the confidence interval of any single land class (regardless of level)
based on
# the summarized polygons of the land use classes. This lets you see how
confident
# the NLCD map believes its assignments to be for any given land cover
class.
# It requires a list of buffered and erased nlcd polygons, the base
confidence raster,
# and an outpath to save the files to, as well as the column keys for the
land use
# classes.
# It returns a dataframe with the summarized values of the confidence
intervals.

def confidence_interval_extract(in_nlcd_polygons, in_conf_raster, outpath,
                               col_keys, grid_or_val="gridcode",
                               delete_yn=False):

    fixed_outpath = filepath_end_slash(outpath)
    outer_cnfinf_df = pd.DataFrame(columns=col_keys)

    for file in in_nlcd_polygons:
        index_name = get_index_name(file)

        print(f"Extracting Confidence Interval Metrics for {index_name}")

        out_table = f"{fixed_outpath}{index_name}_cnfint_table.dbf"
        out_layer = f"{fixed_outpath}{index_name}_cnfint_layer.lyr"

        arcpy.sa.ZonalStatisticsAsTable(
            in_zone_data=file,
            zone_field=grid_or_val,
            in_value_raster=in_conf_raster,
            out_table=out_table,
            ignore_nodata="DATA",
            statistics_type="MEAN_STD",
            process_as_multidimensional="CURRENT_SLICE",
            percentile_values=[90],
            percentile_interpolation_type="AUTO_DETECT",
            circular_calculation="ARITHMETIC",
            circular_wrap_value=360,
            out_join_layer=out_layer
        )

        out_table_list = [out_table]

        cnfint_df = extract_multi_row_to_df(out_table_list, col_keys,
                                             data_extract_field=["MEAN",
                                                                  "STD"],
                                             base_fields=grid_or_val)

        index_list = []

```

```

    for idx in cnfint_df.index:
        idx_new = f"{idx}_cnfint"
        index_list.append(idx_new)

    cnfint_df.index = index_list

    outer_cnfint_df = pd.concat([outer_cnfint_df, cnfint_df])

    if delete_yn is True:
        del_list = [out_table, out_layer]
        delete_from_list(del_list)

    return outer_cnfint_df

'''
# ## 13. STREAMLINING THE IMPLEMENTATION OF ALL THESE FUNCTIONS ## #
'''
# This function combines all the previous functions into one singular,
# streamlined process where you can feed the arguments into this function
# and it will produce its final output of this entire process: a .csv file
# that contains the hectare area, hectare area percent, confidence interval
# mean, and confidence interval std for ALL sites and ALL buffer distances.
# Optionally, it also leaves the step files in-tact so you can see what the
# area around the sites actually looks like.
# It returns a variable dataframe of the final saved .csv file.
# Its variables are as follows: in_sites_file is the shapefile of
site/zones
# polygons to use as the focal point of the surrounding matrix.
# Buffer_dist_list is the list of buffer distances to calculate areas for.
# in_nlcd_polygon is the shapefile of the NLCD polygon to use, usually
variable
# by the level of classification you want to use (level 0 or 2, for
example)
# Scratch_filepath is where to store the step files.
# output_filepath is where to save the final .csv file.
# Site_siterate_list is the list of site numbers used to iterate over the
site polygon.
# landuse_col_keys is the key to the land use classification types.
# coord_sys is the ArcGIS coordinate system object used for this analysis.
# confidence_filepath is the raster of confidence interval NLCD data to
extract
# mean confidence from.
# new_filename is what you want the final .csv to be called.
# erase_siteyn is whether you will erase the sites/zones from the buffer
when
# calculating total area. Recommended to always be TRUE. I can't think of
an
# example where you would want to include the site's land use itself into
your
# analysis of the surrounding area. So keep it TRUE.
# Delete_yn is whether you will delete all stepfiles after analysis.
Defaults to
# false. Recommended set to TRUE because running the analysis subsequent
times
# usually trips up if some files already exist.
# new_col is whether you will add a new column with the total area of the
# analyzed area. Default set to FALSE. though it interacts with confidence
# intervals just fine.

```

```

def streamlined_nlcd_analysis(in_sites_file, buffer_dist_list,
                             in_nlcd_polygon,
                             scratch_filepath, output_filepath,
                             site_iterate_list,
                             landuse_col_keys, coord_sys,
                             confidence_filepath,
                             new_filename="default_filename_complete",
                             erase_siteyn=True, delete_yn=False,
                             new_col=False):

    def timecheck(process_start, function_step):
        print(f"Time elapsed since beginning:
{time_from_start(process_start)}\n"
              f"Time elapsed for this step:
{time_from_start(function_step)}\n")

        functionBeginTime = datetime.datetime.now()
        print(f"Now beginning the streamlined NLCD Analysis on
{str(len(site_iterate_list))} sites.\n")

        # SELECTING SITES AND BUFFERING/ERASING THEM FOR EACH SITE
        bufferingBeginTime = datetime.datetime.now()
        print(f"Now buffering and erasing sites.\n")

        erased_buffers_list = select_and_buffer_many(in_sites_file,
scratch_filepath,
                                                    site_iterate_list,
buff_dist=buffer_dist_list,
                                                    delete_step_yn=delete_yn,
erase_site_yn=erase_siteyn)

        print("Buffering and erasing sites complete.\n")
        timecheck(functionBeginTime, bufferingBeginTime)

        # EXTRACTING THE LAND COVER FOR EACH BUFFER
        extractingBeginTime = datetime.datetime.now()
        print("Extracting land cover data from NLCD files.\n")

        extracted_landcover = land_cover_data_extract(erased_buffers_list,
in_nlcd_polygon,
                                                    scratch_filepath,
coordinate_system=coord_sys)

        print("Land cover data extraction complete.\n")
        timecheck(functionBeginTime, extractingBeginTime)

        # GETTING THE LAND COVER DATA TO A DATAFRAME
        rowsAreaBeginTime = datetime.datetime.now()
        print("Extracting rows to a dataframe and finding the area\n")

        baserows_df = extract_multi_row_to_df(extracted_landcover,
landuse_col_keys)
        baserows_area_df = find_percent_of_area(baserows_df,
new_column_yn=new_col)

        print("Rows extracted to a dataframe and area found.\n")
        timecheck(functionBeginTime, rowsAreaBeginTime)

        # GETTING THE CONFIDENCE INTERVAL INFORMATION
        cnfintBeginTime = datetime.datetime.now()
        print("Extracting confidence interval NLCD data.\n")

```



```

cnfint_df = confidence_interval_extract(extracted_landcover,
confidence_filepath, scratch_filepath,
                                     landuse_col_keys,
delete_yn=delete_yn)

print("Confidence intervals extracted from NLCD data.\n")
timecheck(functionBeginTime, cnfintBeginTime)

# FINALIZING THE DATAFRAME AND SAVING IT
finalizingBeginTime = datetime.datetime.now()
print("Finalizing dataframes and preparing to save file.\n")

output_path =
f"{filepath_end_slash(output_filepath)}{new_filename}.csv"
combined_df = pd.concat([baserows_area_df, cnfint_df])
sorted_combined_df = combined_df.sort_index()
sorted_combined_df.to_csv(output_path)

print(f"Final combined dataframe saved to a csv file.\n"
      f"File saved to: {output_path}\n")
timecheck(functionBeginTime, finalizingBeginTime)

# IF APPLICABLE: DELETING STEP FILES
outfolder_pathform = Path(scratch_filepath)
file_count = sum(1 for file in outfolder_pathform.iterdir() if
file.is_file())

if delete_yn is True:
    deleteBeginTime = datetime.datetime.now()
    print("Now deleting all step files produced in this process.\n")

    delete_from_list(erased_buffers_list)
    delete_from_list(extracted_landcover)

    print("Deleting complete.\n")
    timecheck(functionBeginTime, deleteBeginTime)

# END OF THE FUNCTION
perms = len(site_iterate_list) * len(buffer_dist_list)
varis = perms * len(landuse_col_keys)
print(f"This process is now complete. You have analysed the NLCD
landcover for...\n"
      f"A total of {str(len(site_iterate_list))} sites/zones across
{str(len(buffer_dist_list))}"
      f" buffers and {str(len(landuse_col_keys))} land use classes.\n"
      f"You created at least {file_count} files across {perms} site-
buffer permutations with "
      f"{varis} land use-class variations.\n"
      f"Total time elapsed: {time_from_start(functionBeginTime)}\n"
      f"Have a nice day. I love you.\n")

return sorted_combined_df

'''
# ## 14. Changelog, Backups, and old versions of code ## #
'''

# v1.1 updated runscrip and this one to include options for "level 3"
# data, which is just another kind of classification that divides

```

```
# ecosystems into open and closed, per Anna Treydte's suggestions.
# took around 1hr45min.

# v1.0 First full "release" of the code, first ran with the runsript
# on january 20th 2025, took around 5 hours to complete.
```

Python-Based Data Cleaning and Measuring

This includes the code and functions used to clean the floral survey species data, clean the management survey responses, and calculate response and predictor variables for floral surveys, management surveys, and land cover ratios. This is the code for calculating FQI, management sum, rural and suburban classification, etc. Once again, this does not include the code to run these, only the functions themselves.

```
Used_functions_organizing_data.py
# Andy Sima - Jan-Feb 2025 (Copied June 2 2025)
# Master's Thesis Analysis Work
# Using Python and Pandas to organize field data before going into R
# Functions I have used previously to do some sort of organizational work

# Defined Functions for Cleaning Field Data
# Version 1.6 - Dated 14/03/2025 (March)

# ## LIBRARIES NECESSARY FOR THESE FUNCTIONS ## #
import datetime
import pandas as pd
import os
import statistics
import math
from scipy.stats import linregress
import numpy as np
from nlcd_matrix_analysis_functions import *

# ## DEFINING GLOBAL VARIABLES ## #
'''
Setup variables for the functions that I am using here!
Make sure they are consistent with the organizing_field_data information!
'''

filepaths_list_first = [("raw_data_path",

r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS_LIVE\Stockholm University
Classwork\THESIS WORK\full_data_TOTAL\individual_sites\\"),
                        ("output_path",

r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS_LIVE\Stockholm University
Classwork\THESIS WORK\full_data_TOTAL\test_output\\"),
                        ("species_list",

r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS_LIVE\Stockholm University
Classwork\THESIS WORK\full_data_TOTAL\all_unique_species_list.csv"),
                        ("test_raw_data",

r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS_LIVE\Stockholm University
Classwork\THESIS
WORK\full_data_TOTAL\individual_sites\SpeciesList_sitell_deer_grove_east.xlsx"),
```

```

        ("land_use_binary",
         r"U:\Kursbiblioteket\GE9011\student_files\Sima
Andrew David
(ansi0618)\gis_BIGDATA_files\outputs\sites_level3_full_binary_v1-1_1-22-
25.csv"),
        ("land_use_gradient",
         r"U:\Kursbiblioteket\GE9011\student_files\Sima
Andrew David
(ansi0618)\gis_BIGDATA_files\outputs\zones_level3_full_gradient_v1-1_1-22-
25.csv"),
        ("test_raw_data_path",
         r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS\Stockholm Classes\THESIS
WORK\Site_data_docs\full_data\test_sites\\"),
        ("test_output_path",
         r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS\Stockholm Classes\THESIS
WORK\Site_data_docs\full_data\test_sites\\"),
        ("test_raw_data3",
         r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS\Stockholm Classes\THESIS
WORK\Site_data_docs\full_data\test_raw\\"),
        ("survey_responses",
         r"C:\Users\ansi0618\OneDrive\REAL_DOCUMENTS\Stockholm Classes\THESIS
WORK\Site_data_docs\full_data\survey_responses_consolidated.xlsx"))

filepaths_list = slash_fix_to_dict(filepaths_list_first)

sites_list = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18,
19,
              22, 23, 24, 25, 26, 27, 28, 29, 30]
zones_list = list(range(1, 31))
survey_zones_list = list(range(1, 15)) + list(range(17, 31))
buff_dist_list_sites = [2000]
buff_dist_list_zones = [50, 200, 500, 1000, 1500, 2000, 2500, 3000]
land_use_stats_list = ["MEAN_cnfin", "STD_cnfin", "area_ha",
"area_ha_area"]

classes_key_list = {"L0": [11, 12, 21, 22, 23, 24, 31, 41, 42, 43, 52, 71,
81, 82, 90, 95],
                    "L1": [1, 2, 3, 4, 5, 6, 7, 8, 9, 10],
                    "L2": [1, 2, 3, 4, 5],
                    "L3": [1, 2, 3, 4, 5, 6]
                    }

reclassify_classes_list = {"L0": {11: "Open Water", 12: "Perennial
Ice/Snow", 21: "Developed, Open Space",
                                22: "Developed, Low Intensity", 23:
"Developed, Medium Intensity",
                                24: "Developed, High Intensity", 31:
"Barren Land (Rock/Soil/Clay)",
                                41: "Deciduous Forest", 42: "Evergreen
Forest", 43: "Mixed Forest",
                                52: "Shrub/Scrub", 71:
"Grassland/Herbaceous", 81: "Pasture/Hay",
                                82: "Cultivated Crops", 90: "Woody
Wetlands", 95: "Emergent Herbaceous Wetlands"},
                           "L1": {1: "Water", 2: "OpenLandDev", 3:
"LowDev", 4: "HighDev",

```



```

                    5: "BarrenLand", 6: "ForestEcosystem", 7:
"GrassEcosystem",
                    8: "Field", 9: "Farmland", 10:
"WetEcosystem"},
                    "L2": {1: "Water", 2: "Developed", 3:
"Ecosystem", 4: "Farmland", 5: "OpenField"},
                    "L3": {1: "Water", 2: "Developed", 3:
"ClosedEcosystem", 4: "OpenEcosystem",
                    5: "Farmland", 6: "OpenField"}}

land_cover_dict = {11: {"L0": "Open Water", "L1": [1, "Water"], "L2": [1,
"Water"],
                    "L3": [1, "Water"]},
                    12: {"L0": "Perennial Ice/Snow", "L1": [1, "Water"],
"L2": [1, "Water"],
                    "L3": [1, "Water"]},
                    21: {"L0": "Developed, Open Space", "L1": [2,
"OpenLandDev"], "L2": [2, "Developed"],
                    "L3": [2, "Developed"]},
                    22: {"L0": "Developed, Low Intensity", "L1": [3,
"LowDev"], "L2": [2, "Developed"],
                    "L3": [2, "Developed"]},
                    23: {"L0": "Developed, Medium Intensity", "L1": [3,
"LowDev"], "L2": [2, "Developed"],
                    "L3": [2, "Developed"]},
                    24: {"L0": "Developed, High Intensity", "L1": [4,
"HighDev"], "L2": [2, "Developed"],
                    "L3": [2, "Developed"]},
                    31: {"L0": "Barren Land (Rock/Soil/Clay)", "L1": [5,
"BarrenLand"], "L2": [3, "Ecosystem"],
                    "L3": [4, "OpenEcosystem"]},
                    41: {"L0": "Deciduous Forest", "L1": [6,
"ForestEcosystem"], "L2": [3, "Ecosystem"],
                    "L3": [3, "ClosedEcosystem"]},
                    42: {"L0": "Evergreen Forest", "L1": [6,
"ForestEcosystem"], "L2": [3, "Ecosystem"],
                    "L3": [3, "ClosedEcosystem"]},
                    43: {"L0": "Mixed Forest", "L1": [6, "ForestEcosystem"],
"L2": [3, "Ecosystem"],
                    "L3": [3, "ClosedEcosystem"]},
                    52: {"L0": "Shrub/Scrub", "L1": [7, "GrassEcosystem"],
"L2": [3, "Ecosystem"],
                    "L3": [3, "ClosedEcosystem"]},
                    71: {"L0": "Grassland/Herbaceous", "L1": [7,
"GrassEcosystem"], "L2": [3, "Ecosystem"],
                    "L3": [4, "OpenEcosystem"]},
                    81: {"L0": "Pasture/Hay", "L1": [8, "Field"], "L2": [5,
"OpenField"],
                    "L3": [6, "OpenField"]},
                    82: {"L0": "Cultivated Crops", "L1": [9, "Farmland"],
"L2": [4, "Farmland"],
                    "L3": [5, "Farmland"]},
                    90: {"L0": "Woody Wetlands", "L1": [10, "WetEcosystem"],
"L2": [3, "Ecosystem"],
                    "L3": [3, "ClosedEcosystem"]},
                    95: {"L0": "Emergent Herbaceous Wetlands", "L1": [10,
"WetEcosystem"], "L2": [3, "Ecosystem"],
                    "L3": [4, "OpenEcosystem"]}
}

```

```

rural_suburban_land_class_key = {2: (1, 2, 6, 7, 8, 9, 10, 11, 12, 13, 17,
21, 22, 27, 30),
5: (3, 4, 5, 14, 15, 16, 18, 19, 20, 23,
24, 25, 26, 28, 29)}

quadrat_number = ["r01", "r02", "r03", "r04", "r05", "r06", "r07", "r08",
"r09", "r10"]

quadrat_presence = ["r01_presence", "r02_presence", "r03_presence",
"r04_presence", "r05_presence",
"r06_presence", "r07_presence", "r08_presence",
"r09_presence", "r10_presence", ]

quadrat_coverage = ["r01_coverage", "r02_coverage", "r03_coverage",
"r04_coverage", "r05_coverage",
"r06_coverage", "r07_coverage", "r08_coverage",
"r09_coverage", "r10_coverage", ]

coverage_midpoints = {0: 0, 1: 0.005, 2: 0.03, 3: 0.15, 4: 0.375, 5: 0.625,
6: 0.85, 7: 0.975}

site_metrics = ["Sp. Richness Native. Site", "Sp. Richness Native. Plots
avg.",
"Sp. Richness Native. Plots std.",
"Sp. Richness Introduced. Site", "Sp. Richness Introduced.
Plots avg.",
"Sp. Richness Introduced. Plots std.",
"Sp. Richness All. Site", "Sp. Richness All. Plots avg.",
"Sp. Richness All. Plots std.",
"Percent Introduced Sp. Site", "Percent Introduced Sp.
Plots avg.",
"Percent Introduced Sp. Plots std.",
"Mean C Native. Site", "Mean C Native. Plots avg.", "Mean C
Native. Plots std.",
"Mean C All. Site", "Mean C All. Plots avg.", "Mean C All.
Plots std.",
"Mean C All-Native Diff. Site", "Mean C All-Native Diff.
Plots",
"Median C Native. Site", "Median C Native. Plots avg.",
"Median C Native. Plots std.",
"Median C All. Site", "Median C All. Plots avg.", "Median C
All. Plots std.",
"Median C All-Native Diff. Site", "Median C All-Native
Diff. Plots",
"Mean-based FQI Native. Site", "Mean-based FQI Native.
Plots avg.",
"Mean-based FQI Native. Plots std.", "Mean-based FQI All.
Site",
"Mean-based FQI All. Plots avg.", "Mean-based FQI All.
Plots std.",
"Mean-based FQI All-Native Diff. Site", "Mean-based FQI
All-Native Diff. Plots",
"Mean-based AFQI All. Site", "Mean-based AFQI All. Plots
avg.",
"Mean-based AFQI All. Plots Std.",
"Median-based FQI Native. Site", "Median-based FQI Native.
Plots avg.",
"Median-based FQI Native. Plots std.", "Median-based FQI
All. Site",
"Median-based FQI All. Plots avg.", "Median-based FQI All.
Plots std.",

```

```

        "Median-based FQI All-Native Diff. Site", "Median-based FQI
All-Native Diff. Plots",
        "Median-based AFQI All. Site", "Median-based AFQI All.
Plots avg.",
        "Median-based AFQI All. Plots Std.",
        "Shannon Diversity Index Native. Site", "Shannon Diversity
Index All. Site",
        "Shannon Diversity Index All-Native Diff. Site",
        "Simpson Diversity Index Native. Site", "Simpson Diversity
Index All. Site",
        "Simpson Diversity Index All-Native Diff. Site",
        "Introduced Species Encroachment. Site", "Woody Pest
Encroachment. Site",
        "Herbaceous Pest Encroachment. Site", "Total Pest
Encroachment. Site"
    ]

freq_dict = {"Do not use this technique": 0,
            "Several Times a Year": 5,
            "Once a Year": 4,
            "Every 2-4 Years": 3,
            "Every 5-10 Years": 2,
            "Greater than 10 years": 1}

percent_dict = {"Do not use this technique": 0,
               "Spot-Treat Problem Areas (less than 5% of area managed at
once)": 1,
               "Between 5% and 25%": 2,
               "Between 25% and 50%": 3,
               "Between 50% and 75%": 4,
               "Greater than 75% At a Time": 5}

timeago_dict = {"Do not use this technique": 0,
               "Less than 5 years ago": 1,
               "Between 5 and 10 years ago": 2,
               "Between 10 and 20 years ago": 3,
               "Between 20 and 35 years ago": 4,
               "Between 35 and 50 years ago": 5,
               "Greater than 50 years ago": 6}

permstaff_dict = {"No permanent staff; this site is largely unmanaged": 0,
                 "Between 1 and 3 full-time staff": 1,
                 "between 3 and 5": 2,
                 "between 5 and 7": 3,
                 "between 7 and 10": 4,
                 "between 10 and 15": 5,
                 "between 15 and 20": 6,
                 "between 20 and 30": 7,
                 "between 30 and 50": 8,
                 "more than 50 permanent staff": 9}

seasonstaff_dict = {"No seasonal staff": 0,
                   "Between 1 and 3 seasonal staff": 1,
                   "between 3 and 5": 1.5,
                   "between 5 and 7": 2,
                   "between 7 and 10": 2.5,
                   "between 10 and 15": 3,
                   "between 15 and 20": 3.5,
                   "between 20 and 30": 4,
                   "between 30 and 50": 4.5,
                   "more than 50 seasonal staff": 5}

```



```

volunteer_dict = {"No volunteers": 0,
                  "Between 1 and 3 volunteers": 1,
                  "between 3 and 5": 1.5,
                  "between 5 and 7": 2,
                  "between 7 and 10": 2.5,
                  "between 10 and 15": 3,
                  "between 15 and 20": 3.5,
                  "between 20 and 30": 4,
                  "between 30 and 50": 4.5,
                  "between 50 and 75": 5,
                  "between 75 and 100": 6,
                  "between 100 and 150": 7,
                  "between 150 and 200": 8,
                  "greater than 200 volunteers": 9
                  }

contractor_dict = {"yes": 3,
                  "no": 0}

site_budget_dict = {"Less than $100,000": 50000,
                   "Between $100,000 and $500,000": 300000,
                   "Between $500,000 and $1 million": 750000}

number_other_sites_dict = {"Only this site": 1,
                           "1 to 2 other sites": 2,
                           "3 to 5 other sites": 4,
                           "10 to 15 other sites": 12.5,
                           "5 to 10 other sites": 7.5,
                           "15 to 25 other sites": 20,
                           "60 to 80 other sites": 70,
                           "greater than 100 other sites": 150}

org_budget_dict = {"Between $100,000 and $500,000": 300000,
                  "Between $1 and $2.5 Million": 1750000,
                  "between $2.5 and $5 million": 3750000,
                  "between $15 million and $20 million": 17500000,
                  "between $20 and $30 million": 25000000,
                  "between $50 and $75 million": 62500000,
                  "between $100 and $150 million": 125000000,
                  "between $150 and $200 million": 175000000,
                  "Between $600 million and $1 billion": 800000000,
                  "greater than $1 billion": 1500000000}

dict_of_keydicts = {"freq_dict": freq_dict,
                   "percent_dict": percent_dict,
                   "timeago_dict": timeago_dict,
                   "permstaff_dict": permstaff_dict,
                   "seasonstaff_dict": seasonstaff_dict,
                   "volunteer_dict": volunteer_dict,
                   "contractor_dict": contractor_dict,
                   "site_budget_dict": site_budget_dict,
                   "number_other_sites_dict": number_other_sites_dict,
                   "org_budget_dict": org_budget_dict}

# ## DEFINING FUNCTIONS FOR ORGANIZING FIELD DATA ## #
# ## NOT SITE METRICS ## #
'''
# ## DEFINED FUNCTIONS THAT I HAVE USED PREVIOUSLY ## #
# These are code snippets that were too useful to relegate to just the

```

```

# "code backups" section at the bottom of this document.
'''

# A function that splits up the name of the incoming text and returns
# a list of the fixed names.
def split_up_name(species_name):
    remove_paren = species_name.replace(")", "")
    name_list = remove_paren.split("(")
    sciname = name_list[0][:-1]
    split_name_list = [sciname, name_list[1], species_name]
    return split_name_list

# A function that splits up the name of the incoming text and then adds it
# to a new dataframe.
# It also optionally saves this to a new .csv if you supply a filepath.
def split_up_name_to_df(df_with_names, cols=None, savepath=None):
    if cols is None:
        cols = ["Scientific name", "Common name", "Full name", ]

    split_name_df = pd.DataFrame(columns=cols)

    for i in df_with_names:
        name_list = split_up_name(i)
        split_name_df.loc[len(split_name_df)] = name_list

    if savepath is not None:
        split_name_df.to_csv(savepath)

    return split_name_df

# A function to call all of the raw data spreadsheets and return a list of
# their filepaths.
# Example: filepaths_list["raw_data_path"]
def get_raw_spreadsheets(filepath):
    raw_data_spreadsheets = [os.path.join(filepath, file) for file in
                             os.listdir(filepath)]
    return raw_data_spreadsheets

# A function to produce a list of all the species names from the raw data
# that
# do not match the established species list. Does not fix them,
# just produces an error report with locations.
def get_wrong_names(specieslist, raw_data_folder, outpath,
species_column="Full name data key",
                    outname="wrong_names.csv"):
    species_spreadsheet = pd.read_csv(specieslist)
    species_keys_df = species_spreadsheet[species_column]
    species_keys_list = species_keys_df.tolist()

    checkup_df = pd.DataFrame(columns=["wrong name", "filepath"])
    good_count = 0
    bad_count = 0

    raw_data = get_raw_spreadsheets(raw_data_folder)

    for i in raw_data:
        spread_df = pd.read_excel(i)

```

```

species_df = spread_df.iloc[1:, 0]
strip_i = os.path.basename(i)

for j in species_df:
    if j not in species_keys_list:
        print("Bad name!")
        checkout_df.loc[len(checkout_df)] = [j, strip_i]
        bad_count += 1
    else:
        print("Good name!")
        good_count += 1

if bad_count != 0:
    new_file_name = str(outpath) + outname
    checkout_df.to_csv(new_file_name)
    print(f"Bad names: {bad_count}.")
else:
    print("No bad names!")

print(f"Good names: {good_count}.")

return checkout_df

# function for zipping site names into an iterable list
# this is bad and only works for binary
def binary_zipped_sites(sites_list_i, ):
    rows_perc_iterate_list_i = []
    rows_area_iterate_list_i = []
    site_num_list_i = []
    for i_site in sites_list_i:
        iter_i = f"site{i_site}_buff2000m_LEVEL3_area_ha_%area"
        name_i = f"site{i_site}_landuse_class"
        area_i = f"site{i_site}_buff2000m_LEVEL3_area_ha"
        rows_perc_iterate_list_i.append(iter_i)
        site_num_list_i.append(name_i)
        rows_area_iterate_list_i.append(area_i)

    zipped_sites_i = list(zip(site_num_list_i, rows_perc_iterate_list_i,
rows_area_iterate_list_i))
    return zipped_sites_i

# A function useful for analyzing landscape and land use data for sites
# and for zones. Currently this only works for binary comparison data.
# trying to get landscape cover data to present itself nicely
# the zipped iterator MUST have three strings in this order:
# The name of row in the final dataset,
# the name of the area percent row in the input dataset,
# the name of the area value row in the input dataset.
def get_binary_land_use(filepath, outpath, sites_list_i,
    binary_classes_list=None, outname=None):
    if binary_classes_list is None:
        binary_classes_list = ["land_class", "class_number",
"area_r_over_s_ratio", "area_val_r", "area_val_s",
"perc_val_r", "perc_val_s", "total_area",
"max_class"]

    if outname is None:
        outname = "NEW_landuse_sites_level3_binary_classes_1-30-25.csv"

```



```

binary_data = pd.read_csv(filepath, index_col=0)

binary_classes = pd.DataFrame(columns=binary_classes_list)

zipped_iterator = binary_zipped_sites(sites_list_i)

for f, j, k in zipped_iterator:
    rural_or_sub = binary_data.loc[j, "5"] > binary_data.loc[j, "2"]
    max_class = binary_data.loc[j,].drop(["total_area"]).idxmax()
    total_area = binary_data.loc[j, "total_area"]

    ratio_area = (binary_data.loc[k, "5"] / binary_data.loc[k, "2"])
    area_r = binary_data.loc[k, "5"]
    area_s = binary_data.loc[k, "2"]

    percent_r = binary_data.loc[j, "5"]
    percent_s = binary_data.loc[j, "2"]

    if rural_or_sub:
        site_data = ["Rural", "5", ratio_area, area_r, area_s,
percent_r, percent_s, total_area, str(max_class)]
    else:
        site_data = ["Suburban", "2", ratio_area, area_r, area_s,
percent_r, percent_s, total_area, str(max_class)]

    binary_classes.loc[f] = site_data

    save_file = outpath + outname
    binary_classes.to_csv(save_file)
    print(f"File saved to: {save_file}")

    return binary_classes

def summary_values_suite(data, row_name, dom_class,
                        total_area="total_area", ):
    total_area = data.loc[row_name, total_area]

    dataset = data[["1", "2", "3", "4", "5", "6"]]

    max_class = dataset.loc[row_name].idxmax()

    dominant_rs_class = dataset[["2", "5"]].loc[row_name].idxmax()

    binary_dominant_class = dom_class

    dominant_wilderness = dataset[["3", "4"]].loc[row_name].idxmax()

    return max_class, total_area, dominant_rs_class, binary_dominant_class,
dominant_wilderness

def percent_and_area_values_suite(data, perc_row, ):
    # this works for both area and percent values, you just
    # need to change what the input row is, area or percent

    percent_total_wild = data.loc[perc_row, "3"] + data.loc[perc_row, "4"]
    percent_open_wild = data.loc[perc_row, "4"]
    percent_closed_wild = data.loc[perc_row, "3"]
    percent_rural = data.loc[perc_row, "5"]
    percent_suburban = data.loc[perc_row, "2"]

```

```

percent_ignored_field = data.loc[perc_row, "6"]
return (percent_total_wild, percent_open_wild, percent_closed_wild,
percent_rural,
percent_suburban, percent_ignored_field)

def ratios_values_suite(data, perc_row, bi_dom_class):
total_wild = data.loc[perc_row, "3"] + data.loc[perc_row, "4"]
dom_class = data.loc[perc_row, bi_dom_class]
closed_wild = data.loc[perc_row, "3"]
suburban = data.loc[perc_row, "2"]

if dom_class == 0:
ratio_tw_dom = np.nan
ratio_ow_dom = np.nan
ratio_cw_dom = np.nan
else:
ratio_tw_dom = total_wild / dom_class
ratio_ow_dom = data.loc[perc_row, "4"] / dom_class
ratio_cw_dom = data.loc[perc_row, "3"] / dom_class

if closed_wild == 0:
ratio_ow_cw = np.nan
else:
ratio_ow_cw = data.loc[perc_row, "4"] / closed_wild

if suburban == 0:
ratio_r_s = np.nan
else:
ratio_r_s = data.loc[perc_row, "5"] / suburban

return ratio_tw_dom, ratio_ow_dom, ratio_cw_dom, ratio_ow_cw, ratio_r_s

def zone_summary_values(zone_data, buff_list, zone_num, summary_vals_list,
summary_dict_tag_list=None, min_data_value=4):
if summary_dict_tag_list is None:
summary_dict_tag_list = ["avg_", "stdev_", "slope_", "intercept_",
"r_squared_", "relative_slope_", ]

zone_columns = zone_data.columns.values.tolist()
row_name = f"zone_{zone_num}"

zone_summary_big = pd.DataFrame()

for summary_val in summary_vals_list:
filter_columns = [col for col in zone_columns if summary_val in
col]

filtered_row = zone_data.loc[row_name, filter_columns]
filtered_row_vals = filtered_row.tolist()

buff_list_x = np.array(buff_list)
filter_list_y = np.array(filtered_row_vals)

nan_mask = ~np.isnan(buff_list_x) & ~np.isnan(filter_list_y)

buff_list_numpy = buff_list_x[nan_mask]
filtered_row_numpy = filter_list_y[nan_mask]

if len(buff_list_numpy) < min_data_value:
print(len(buff_list_numpy))

```

```

        summary_avg, summary_std, slope, intercept, r_val, r_val2,
p_val, std_err, = np.nan, np.nan, np.nan, np.nan, np.nan, np.nan, np.nan,
np.nan
    else:
        summary_avg = filtered_row_numpy.mean()
        summary_std = filtered_row_numpy.std()
        slope, intercept, r_val, p_val, std_err =
linregress(buff_list_numpy.astype(float),
filtered_row_numpy.astype(float))
        r_val2 = r_val ** 2

    if summary_avg == 0 or summary_avg is np.nan:
        rel_slope = np.nan
    else:
        rel_slope = slope / summary_avg

    summary_metrics = (summary_avg, summary_std, slope, intercept,
r_val2, rel_slope)
    summary_columns = [item + summary_val for item in
summary_dict_tag_list]

    summary_val_df = pd.DataFrame(columns=summary_columns)
    summary_val_df.loc[row_name] = summary_metrics

    zone_summary_big = pd.concat([zone_summary_big, summary_val_df],
axis=1)

    return zone_summary_big

# get_gradient_land_use(filepaths_list["land_use_gradient"],
filepaths_list["test_output_path"],
#
#         zones_list, rural_suburban_land_class_key,
buff_dist_list_zones, replace_zero=False,
#
outname="landuse_zones_l3_gradient_summary_test_zero5.csv")
def get_gradient_land_use(raw_filepath, outpath, zone_list_iterator,
binary_land_class_key,
                        buffer_list, base_classes_list_buffer=None,
base_classes_list_summary=None,
                        land_metrics_tag=None, total_area="total_area",
                        outname="landuse_zones_l3_gradient_summary.csv",
buffer_tag="_buff{ }m",
                        replace_zero=False, replace_val=0.01,
min_data_value=4,
                        replace_data_tag="err"):
    if base_classes_list_buffer is None:
        base_classes_list_buffer = {"summary_values": ["max_class",
"total_area", "dominant_rs_class",
"binary_dominant_class", "dominant_wilderness", ],
                                "percents": ["percent_total_wild",
"percent_open_wild", "percent_closed_wild",
                                "percent_rural",
"percent_suburban", "percent_ignored_field", ],
                                "areas": ["area_total_wild",
"area_open_wild", "area_closed_wild",
                                "area_rural",
"area_suburban", "area_ignored_field"],

```



```

        "ratio_openwild_dominant",
        "ratio_closedwild_dominant",
        "ratio_openwild_closedwild",
        "ratio_rural_suburban", ], }

    if base_classes_list_summary is None:
        base_classes_list_summary = (base_classes_list_buffer["percents"] +
                                     base_classes_list_buffer["areas"] +
                                     base_classes_list_buffer["ratios"])

    if land_metrics_tag is None:
        land_metrics_tag = ["MEAN_cnfin", "STD_cnfin", "area_ha",
                             "area_ha_%area"]

    process_start = datetime.datetime.now()
    print("Beginning to extract gradient land cover data now.\n")

    gradient_data = pd.read_csv(raw_filepath, index_col=0, header=0)
    if replace_zero is True:
        gradient_data.replace(0, replace_val, inplace=True)

    gradient_data_results = pd.DataFrame()

    for zone in zone_list_iterator:
        print(f"Extracting gradient data for zone {zone}.")
        if zone in binary_land_class_key[2]:
            dominant_class = "2"
            secondary_class = "5"
        else:
            dominant_class = "5"
            secondary_class = "2"

        base_classes_buffer_iterator = [item for sublist_1 in
                                       base_classes_list_buffer.values()
                                       for item in sublist_1]

        zone_base_classes_dataframe = pd.DataFrame()

        for buffer in buffer_list:
            base_tag_i = f"site{zone}_buff{buffer}_m_LEVEL3_"
            results_buff_tag_i = buffer_tag.format(buffer)
            col_list = [tag + results_buff_tag_i for tag in
                        base_classes_buffer_iterator]
            row_name = f"zone_{zone}"

            area_row = base_tag_i + land_metrics_tag[2]
            perc_row = base_tag_i + land_metrics_tag[3]

            summary_values = summary_values_suite(gradient_data, area_row,
            dominant_class)
            percents_values = percent_and_area_values_suite(gradient_data,
            perc_row)
            area_values = percent_and_area_values_suite(gradient_data,
            area_row)
            ratio_values = ratios_values_suite(gradient_data, area_row,
            dominant_class)

            buffer_values = (summary_values + percents_values + area_values
            + ratio_values)
            buff_iter_vals_df = pd.DataFrame(columns=col_list)

```

```

        buff_iter_vals_df.loc[row_name] = buffer_values

        zone_base_classes_dataframe =
pd.concat([zone_base_classes_dataframe, buff_iter_vals_df], axis=1)

        zone_summary = zone_summary_values(zone_base_classes_dataframe,
buffer_list,
                                                zone, base_classes_list_summary,
min_data_value=min_data_value)

        full_zone_data = pd.concat([zone_base_classes_dataframe,
zone_summary], axis=1)
        gradient_data_results = pd.concat([gradient_data_results,
full_zone_data])

        gradient_data_results.replace(np.nan, replace_data_tag, inplace=True)

        print("All zone data extracted.\n")

        outfile = outpath + outname
        print(f"Saving file to: \n{outfile}\n")

        gradient_data_results.to_csv(outfile)

        process_endtime = datetime.datetime.now()
        print("Gradient data extraction complete.")
        print(f"Time elapsed: {str(process_endtime - process_start)}")
        print("Have a great day. I love you.")

        return gradient_data_results

# ## DEFINING FUNCTIONS FOR ORGANIZING FIELD DATA ## #
# ## FUNCTIONS TO PRODUCE SITE METRICS ## #
'''
These are the functions are will actually produce all the metrics that I
want
for each given site! They are very important! Each one produces one part of
a
the list of different metrics.
'''

# presence iterator only
def generic_quads_wrapper(dataset, func, iterator_inner=None, kwargs=None,
step_list=False):
    if iterator_inner is None:
        iterator_inner = quadrat_presence

    generic_quads_list = []

    for j in iterator_inner:
        quad_val = func(dataset, j, kwargs)
        generic_quads_list.append(quad_val)

    #print(generic_quads_list)

    generic_quads_avg = statistics.mean(generic_quads_list)
    generic_quads_std = statistics.stdev(generic_quads_list)

    if step_list is True:

```

```

        return generic_quads_list
    else:
        return generic_quads_avg, generic_quads_std

def richness_quads(dataset, iterative, kwargs=None, ):
    sumval = float(dataset[iterative].sum())
    return sumval

def percent_introduced_quads(dataset, iterative, kwargs=None, ):
    if kwargs is None:
        kwargs = "Native Plant"

    introduced_val = float(dataset[iterative][dataset[kwargs].isin([0,
2])]
.sum())
    sumval = float(dataset[iterative].sum())
    percent_int_val = (introduced_val / sumval) * 100

    return percent_int_val

def mean_c_quads(dataset, iterative, kwargs=None, ):
    if kwargs is None:
        cval_col = "CoC"
    else:
        cval_col = kwargs[0]

    present_species_cvals = dataset[cval_col][dataset[iterative] ==
1].astype(float)

    if len(present_species_cvals) == 0 or present_species_cvals is None:
        c_mean = 0
    else:
        c_mean = present_species_cvals.mean()

    return c_mean

def med_c_quads(dataset, iterative, kwargs=None, ):
    if kwargs is None:
        kwargs = "CoC"

    present_species_cvals = dataset[kwargs][dataset[iterative] ==
1].astype(float)

    if len(present_species_cvals) == 0 or present_species_cvals is None:
        c_med = 0
    else:
        c_med = present_species_cvals.median()

    return c_med

def fqi_quads(all_dataset, iterative,
               kwargs=None, ):
    if kwargs is None:
        nat_col = "Native Plant"
        nat = True
        mean_med_func = mean_c_quads
    else:

```



```

        nat_col = kwargs[0]
        nat = kwargs[1]
        mean_med_func = kwargs[2]

    only_nat_sp = all_dataset[all_dataset[nat_col] == 1]
    n_sumval = float(only_nat_sp[iterative].sum())
    sqrt_n = math.sqrt(n_sumval)

    if nat is True:
        mean_med_c = mean_med_func(only_nat_sp, iterative, kwargs=None)
    else:
        all_sp = all_dataset[all_dataset[nat_col] != 3]
        mean_med_c = mean_med_func(all_sp, iterative, kwargs=None)

    fqi_val = mean_med_c * sqrt_n
    return fqi_val

def afqi_func(mean_c, n_nat, n_sp):
    sqrt_nat = math.sqrt(n_nat)
    sqrt_sp = math.sqrt(n_sp)
    return ((mean_c / 10) * (sqrt_nat / sqrt_sp)) * 100

def afqi_quads(all_dataset, iterative,
               kwargs=None, ):
    if kwargs is None:
        col = "Native Plant"
        mean_med_func = mean_c_quads
    else:
        col = kwargs[0]
        mean_med_func = kwargs[1]

    only_nat_sp = all_dataset[all_dataset[col] == 1]
    n_sumval = float(only_nat_sp[iterative].sum())
    s_sumval = all_dataset[iterative].sum()

    mean_med_c_nat = mean_med_func(only_nat_sp, iterative)
    afqi_val = afqi_func(mean_med_c_nat, n_sumval, s_sumval)

    return afqi_val

# Presence iterator only
def species_richness_and_perc_intro(native_species, all_species,
iterator=None,
                                   native_col="Native Plant"):
    if iterator is None:
        iterator = quadrat_presence

    richness_list = [len(native_species)]
    richness_list.extend(generic_quads_wrapper(native_species,
richness_quads, iterator_inner=iterator))

    introduced_species = all_species[all_species[native_col].isin([0, 2])]
    richness_list.append(len(introduced_species))
    richness_list.extend(generic_quads_wrapper(introduced_species,
richness_quads, iterator_inner=iterator))

    richness_list.append(len(all_species))

```

```

    richness_list.extend(generic_quads_wrapper(all_species, richness_quads,
iterator_inner=iterator))

    richness_list.append((len(introduced_species) / len(all_species)) *
100)
    richness_list.extend(generic_quads_wrapper(all_species,
percent_introduced_quads, iterator_inner=iterator))

    return richness_list

# Presence iterator only
def mean_c_med_c_and_diff(native_species, all_species, iterator=None,
                           cval_col="CoC"):
    if iterator is None:
        iterator = quadrat_presence

    def c_on_repeat(func, aux_func):
        c_list = []

        c_nat = aux_func(native_species[cval_col])
        c_nat_quads = generic_quads_wrapper(native_species, func,
iterator_inner=iterator)
        c_all = aux_func(all_species[cval_col])
        c_all_quads = generic_quads_wrapper(all_species, func,
iterator_inner=iterator)

        c_diff_site = c_all - c_nat
        c_diff_quads = c_all_quads[0] - c_nat_quads[0]
        diff_list = (c_diff_site, c_diff_quads)

        c_list.append(c_nat)
        c_list.extend(c_nat_quads)
        c_list.append(c_all)
        c_list.extend(c_all_quads)
        c_list.extend(diff_list)

    return c_list

mean_c_list = c_on_repeat(mean_c_quads, pd.Series.mean)
med_c_list = c_on_repeat(med_c_quads, pd.Series.median)
all_c_list = mean_c_list + med_c_list

return all_c_list

# Presence iterator only
def fqi_afqi_and_diff(all_species_w_unid, native_species, all_species,
iterator=None,
                       cval_col="CoC", native_col="Native Plant",
mean_med_func=mean_c_quads):
    if iterator is None:
        iterator = quadrat_presence

    if mean_med_func == mean_c_quads:
        aux_func = pd.Series.mean
    else:
        aux_func = pd.Series.median

    fqi_afqi_list = []

```

```

mean_med_c_nat_site = aux_func(native_species[cval_col])
sqrt_n_nat_site = math.sqrt(len(native_species))

# here
fqi_nat_site = mean_med_c_nat_site * sqrt_n_nat_site
fqi_afqi_list.append(fqi_nat_site)

# here
fqi_nat_quads = generic_quads_wrapper(all_species_w_unid, fqi_quads,
                                     kwargs=["Native Plant", True,
mean_med_func],
                                     iterator_inner=iterator)

fqi_afqi_list.extend(fqi_nat_quads)

mean_med_c_all_site = aux_func(all_species[cval_col])

# here
fqi_all_site = mean_med_c_all_site * sqrt_n_nat_site
fqi_afqi_list.append(fqi_all_site)

# here
fqi_all_quads = generic_quads_wrapper(all_species_w_unid, fqi_quads,
                                     kwargs=["Native Plant", False,
mean_med_func],
                                     iterator_inner=iterator)

fqi_afqi_list.extend(fqi_all_quads)

# here
fqi_diff_site = fqi_all_site - fqi_nat_site
fqi_diff_quads = fqi_all_quads[0] - fqi_nat_quads[0]
fqi_diff_total = (fqi_diff_site, fqi_diff_quads)
fqi_afqi_list.extend(fqi_diff_total)

# here
afqi_all_site = afqi_func(mean_med_c_nat_site, len(native_species),
len(all_species_w_unid))
fqi_afqi_list.append(afqi_all_site)

# here
afqi_all_plots = generic_quads_wrapper(all_species_w_unid, afqi_quads,
                                     kwargs=["Native Plant",
mean_med_func],
                                     iterator_inner=iterator)

fqi_afqi_list.extend(afqi_all_plots)

return fqi_afqi_list

# Coverage iterator only
def mean_cover_individual(species_row, coverage_iterator=None,
midpoint_dict=None, ):
    if coverage_iterator is None:
        coverage_iterator = quadrat_coverage
    if midpoint_dict is None:
        midpoint_dict = coverage_midpoints

    indiv_coverage_list = []

    for cov in coverage_iterator:
        indiv_coverage_list.append(species_row[cov])

```



```

    cover_dict = {int(cov_class): indiv_coverage_list.count(cov_class) for
cov_class in indiv_coverage_list}

    cover_sum = 0
    pres_sum = len(coverage_iterator)
    for cov_class, cov_total in cover_dict.items():
        cover_sum += float(cov_total) * midpoint_dict[cov_class]
    mean_cover_val = cover_sum / pres_sum

    return mean_cover_val

def mean_cover_datasheet(all_species_w_un, species_dict, iterator):
    full_index_names = species_dict
    species_coverage_dict = {}

    for spec in full_index_names:
        species_row = all_species_w_un.loc[spec]
        spec_val = mean_cover_individual(species_row,
coverage_iterator=iterator)
        species_coverage_dict[spec] = spec_val

    return species_coverage_dict

def relative_cover(mean_coverage_dict):
    total_mc = sum(mean_coverage_dict.values())
    rc_dict = {}

    for spec, mc_val in mean_coverage_dict.items():
        rc_dict[spec] = mc_val / total_mc

    return rc_dict

# presence iterator only
def relative_frequency(all_species_w_un, species_dict,
presence_iterator=None):
    if presence_iterator is None:
        presence_iterator = quadrat_presence

    frequency_dict = {}
    for spec in species_dict:
        spec_sum = 0
        for presence in presence_iterator:
            spec_sum += all_species_w_un.loc[spec, presence]
        frequency_dict[spec] = float(spec_sum)

    frequency_sum = sum(frequency_dict.values())

    rf_dict = {}

    for spec in species_dict:
        rf_dict[spec] = frequency_dict[spec] / frequency_sum

    return rf_dict

def relative_ivi_values(r_cov_dict, r_freq_dict, species_index):
    ivi_val_dict = {}
    for spec in species_index:

```

```

    ivi_val_dict[spec] = r_cov_dict[spec] + r_freq_dict[spec]

ivi_sum = sum(ivi_val_dict.values())

rivi_dict = {}
for spec in species_index:
    rivi_dict[spec] = ivi_val_dict[spec] / ivi_sum

return rivi_dict

def shannon_diversity(rel_ivi_dict, species_index):
    shannon = 0
    for spec in species_index:
        indiv_val = rel_ivi_dict[spec]
        shannon += (indiv_val * math.log(indiv_val))

    true_shannon = -1 * shannon
    return true_shannon

def simpson_diversity(rel_ivi_dict, species_index):
    doh = 0

    for spec in species_index:
        indiv_val = rel_ivi_dict[spec]
        doh += indiv_val ** 2

    true_doh = 1 - doh
    return true_doh

def top_rivi_species(rel_ivi_dict, top_val=5):
    rivi_sorted_dict = {spec: rivi for spec, rivi in
sorted(rel_ivi_dict.items(),
                                                    key=lambda
item: item[1], reverse=True)}
    top_spec = list(list(rivi_sorted_dict.items())[:top_val])
    return top_spec

def top_c_values(dataset, c_val_col="CoC", name_col="Scientific name",
top_val=5):
    dataset_sorted = dataset.sort_values([c_val_col, name_col],
ascending=False).iloc[:top_val, ]
    top_c_list = dataset_sorted.index.values.tolist()
    top_c_dict = {}
    for spec in top_c_list:
        top_c_dict[spec] = dataset_sorted.loc[spec, c_val_col]
    top_c_spec = list(top_c_dict.items())
    return top_c_spec

# This is a supplementary function that takes generic coverage information
and
# figured out what percentage of the total it is.
def encroach_me(mc_tot, species_names, mean_cover_dict):
    encroach_val = 0
    for spec in species_names:
        encroach_val += mean_cover_dict[spec]
    encroach_tot = encroach_val / mc_tot

```

```

    return encroach_tot

# This encroachment metrics function takes the mean cover dictionary of all
# species present
# and calculates the different encroachment metrics for the entire zone
# based on all
# presence and coverages across all plots. It is entirely a zone-wide
# metric.
def encroachment_metrics(mean_cover_dict, dataset,
                          wood_col="Woody Pest", herb_col="Herbaceous Pest",
                          nat_col="Native Plant"):
    total_mc = sum(mean_cover_dict.values())

    woody_species = dataset[dataset[wood_col] == 1].index.values.tolist()
    herb_species = dataset[dataset[herb_col] == 1].index.values.tolist()
    introduced_species = dataset[dataset[nat_col].isin([0,
2])].index.values.tolist()

    introduced_encroach = encroach_me(total_mc, introduced_species,
mean_cover_dict)
    woody_encroach = encroach_me(total_mc, woody_species, mean_cover_dict)
    herby_encroach = encroach_me(total_mc, herb_species, mean_cover_dict)
    total_pest_encroach = woody_encroach + herby_encroach

    encroach_list = (introduced_encroach, woody_encroach, herby_encroach,
total_pest_encroach)
    return encroach_list

# This modified version of encroachment metrics only calculates the
# encroachment for an
# individual plot within an entire zone, making it more suited for mixed
# model metrics.
def modified_plot_encroachment_metrics(data_frame, iterator, midpoint_dict,
native_col="Native Plant",
                                     wood_col="Woody Pest",
herb_col="Herbaceous Pest", nat_col="Native Plant"):
    cover_total = 0
    for row in data_frame[iterator]:
        cover_total += midpoint_dict[row]

    dataset = data_frame[data_frame[nat_col] != 3]
    row_species = dataset[dataset[iterator] != 0].index.values.tolist()
    species_dict = {}

    for species in row_species:
        cover_val = dataset.loc[species, iterator]
        cover_percent = midpoint_dict[cover_val]
        species_dict[species] = cover_percent

    data_this_roi = dataset[dataset[iterator] != 0]
    introduced_species = data_this_roi[data_this_roi[nat_col].isin([0,
2])].index.values.tolist()
    woody_species = data_this_roi[data_this_roi[wood_col] ==
1].index.values.tolist()
    herby_species = data_this_roi[data_this_roi[herb_col] ==
1].index.values.tolist()

    introduced_encroach = encroach_me(cover_total, introduced_species,
species_dict)

```



```

woody_encroach = encroach_me(cover_total, woody_species, species_dict)
herby_encroach = encroach_me(cover_total, herby_species, species_dict)
total_pest_encroach = woody_encroach + herby_encroach

encroach_list = (introduced_encroach, woody_encroach, herby_encroach,
total_pest_encroach)
return encroach_list

# Coverage iterator only
def mean_cover_shannon_shared_metrics(all_data_w_unid, native_species,
coverage_iterator, presence_iterator,
                                     top_val=5, print_extra=False,
outpath=None, ):
    species_index_list = all_data_w_unid.index.values.tolist()
    nat_species_index_list = native_species.index.values.tolist()

    mc_shared_metrics_list = []

    def shannon_simpson_run(dataset, species_index, raw_data, nat_or_all):
        mc_dict = mean_cover_datasheet(dataset, species_index,
iterator=coverage_iterator)
        rc_dict = relative_cover(mc_dict)
        rf_dict = relative_frequency(dataset, species_index,
presence_iterator=presence_iterator)
        rivi_dict = relative_ivi_values(rc_dict, rf_dict, species_index)
        shannon_val = shannon_diversity(rivi_dict, species_index)
        simpson_val = simpson_diversity(rivi_dict, species_index)
        top_species = top_rivi_species(rivi_dict, top_val=top_val)

        if print_extra is True:
            species_index_temp = raw_data
            outfile = outpath +
f"print_extra_rivi_dataframe_{nat_or_all}.csv"
            for dict_i, name_i in ((mc_dict, "mc_dict"), (rc_dict,
"rc_dict"),
                                (rf_dict, "rf_dict"), (rivi_dict,
"rivi_dict")):
                temp_frame = pd.DataFrame.from_dict(dict_i, orient="index",
columns=[name_i])
                species_index_temp = pd.merge(species_index_temp,
temp_frame, left_index=True, right_index=True,
                                                how="left")
                species_index_temp.to_csv(outfile)

        return shannon_val, simpson_val, top_species

    nat_mean_cover = shannon_simpson_run(native_species,
nat_species_index_list, native_species, nat_or_all="native")
    all_mean_cover = shannon_simpson_run(all_data_w_unid,
species_index_list, all_data_w_unid, nat_or_all="all")

    mc_dict_outer = mean_cover_datasheet(all_data_w_unid,
species_index_list, iterator=coverage_iterator)
    encroach_all = encroachment_metrics(mc_dict_outer, all_data_w_unid)

    shannon_all_nat_diff = all_mean_cover[0] - nat_mean_cover[0]
    simpson_all_nat_diff = all_mean_cover[1] - nat_mean_cover[1]

    mc_shared_metrics_list.append(nat_mean_cover[0])
    mc_shared_metrics_list.append(all_mean_cover[0])

```

```

mc_shared_metrics_list.append(shannon_all_nat_diff)
mc_shared_metrics_list.append(nat_mean_cover[1])
mc_shared_metrics_list.append(all_mean_cover[1])
mc_shared_metrics_list.append(simpson_all_nat_diff)
mc_shared_metrics_list.extend(encroach_all)

top_c_species = top_c_values(native_species, top_val=top_val)

top_species_dict = {"Top IVI Species. Native": nat_mean_cover[2],
                    "Top IVI Species. All": all_mean_cover[2],
                    "Top C Species. Native": top_c_species}

return mc_shared_metrics_list, top_species_dict

def combined_metrics_run_single_sheet(merged_data, top_val,
presence_iterator,
coverage_iterator, print_extra,
outpath=None):
    print("Running metrics on the site...")

    full_data_sheet_all_with_unidentified = merged_data

    full_data_sheet_native = full_data_sheet_all_with_unidentified[
        full_data_sheet_all_with_unidentified["Native Plant"] == 1]
    full_data_sheet_all = full_data_sheet_all_with_unidentified[
        full_data_sheet_all_with_unidentified["Native Plant"] != 3]

    metrics_pt1 = species_richness_and_perc_intro(full_data_sheet_native,
full_data_sheet_all_with_unidentified,

iterator=presence_iterator)
    metrics_pt2 = mean_c_med_c_and_diff(full_data_sheet_native,
full_data_sheet_all,

iterator=presence_iterator)
    metrics_pt3 = fqi_afqi_and_diff(full_data_sheet_all_with_unidentified,
full_data_sheet_native,

full_data_sheet_all,

mean_med_func=mean_c_quads,

iterator=presence_iterator)
    metrics_pt4 = fqi_afqi_and_diff(full_data_sheet_all_with_unidentified,
full_data_sheet_native,

full_data_sheet_all,

mean_med_func=med_c_quads,

iterator=presence_iterator)
    metrics_pt5 =
mean_cover_shannon_shared_metrics(full_data_sheet_all_with_unidentified,
full_data_sheet_native,

coverage_iterator=coverage_iterator,

presence_iterator=presence_iterator,

top_val=top_val,

print_extra=print_extra,

outpath=outpath, )

    metrics_total_step = metrics_pt1 + metrics_pt2 + metrics_pt3 +
metrics_pt4 + metrics_pt5[0]

    return metrics_total_step, metrics_pt5[1]

```

```

# full_data_metrics(filepaths_list["raw_data_path"],
filepaths_list["species_list"],
# filepaths_list["output_path"],)
def full_data_metrics(raw_data_folder, species_list, outpath,
metric_names=None,
                        top_species_names=None,
metric_outname="all_sites_full_metrics.csv",
                        top_species_outname="all_sites_top_species.csv",
                        round_numbers=False, round_val=4, top_val=5,
                        presence_iterator=None, coverage_iterator=None,
print_extra=False,
                        save_merge=False):
    if coverage_iterator is None:
        coverage_iterator = quadrat_coverage
    if presence_iterator is None:
        presence_iterator = quadrat_presence
    if metric_names is None:
        metric_names = site_metrics
    if top_species_names is None:
        top_species_names = ["Top IVI Species. Native", "Top IVI Species.
All",
                            "Top C Species. Native"]

    process_start = datetime.datetime.now()

    filelist = get_raw_spreadsheets(raw_data_folder)
    species_sheet = pd.read_csv(species_list, header=0, index_col=0)

    print("Now beginning to measure site metrics...")
    print(f"{len(metric_names)} metrics will be calculated for
{len(filelist)} sites.\n")

    big_richness_df = pd.DataFrame(columns=metric_names)
    big_top_species_df = pd.DataFrame(columns=top_species_names)

    for file in filelist:
        iterate_start = datetime.datetime.now()

        filename = os.path.basename(file)
        site_num_whole = get_site_number(filename, tag1="site",
tag2="Site", delim="_")
        filename_start = filename.split(".")[0].split("_")[1:]
        bald_filename = "_".join(filename_start)

        print(f"Now getting metrics for {site_num_whole}...")

        data_sheet = pd.read_excel(file, header=0, skiprows=[0],
index_col=0)
        data_sheet = data_sheet.fillna(0).astype(int)

        merged_data = pd.merge(data_sheet, species_sheet, left_index=True,
right_index=True, how="left")
        if save_merge is True:
            mergesavename = outpath + site_num_whole + "merged.csv"
            merged_data.to_csv(mergesavename)

        metrics_run = combined_metrics_run_single_sheet(merged_data,
top_val,

```



```

presence_iterator=presence_iterator,

coverage_iterator=coverage_iterator,

outpath=outpath,

print_extra=print_extra)
    metrics_total_step = metrics_run[0]
    metrics_top_species = metrics_run[1]

    if round_numbers is True:
        metrics_total = []
        for met in metrics_total_step:
            r = round(met, round_val)
            metrics_total.append(r)
    else:
        metrics_total = metrics_total_step

    big_richness_df.loc[bald_filename] = metrics_total

    top_species = metrics_top_species
    big_top_species_df.loc[bald_filename] = top_species.values()

    #DEBUG_save_path = outpath +
f"debug_test_{site_num_whole}_file.csv"
    #full_data_sheet_all_with_unidentified.to_csv(DEBUG_save_path)

    iterate_end = datetime.datetime.now()

    print(f"Metrics processed for {site_num_whole} in {str(iterate_end
- iterate_start)}.")
    print(f"Total time elapsed: {str(iterate_end - process_start)}\n")

    metric_save_path = outpath + metric_outname
    top_species_save_path = outpath + top_species_outname

    big_richness_df.to_csv(metric_save_path)
    big_top_species_df.to_csv(top_species_save_path)

    process_end = datetime.datetime.now()

    print(f"All metrics completed. File saved to:\n{metric_save_path}\n")
    print(f"A total of {str((len(metric_names) * len(filelist) +
(len(filelist) * 3)))} metrics have been measured.")
    print(f"Total time elapsed: {str(process_end - process_start)}")
    print("Have a nice day. I love you.")

    return big_richness_df, big_top_species_df, metric_save_path,
top_species_save_path

# This versio of the code does the same thing as full_data_metrics BUT it
runs it on
# the entire dataset (all sites, all observations) as a whole and produces
a
# statistically-worthless spreadsheet which demonstrates the values of
# the entire dataset if it were one site.
# all_sites_combo(filepaths_list["raw_data_path"],
filepaths_list["species_list"],
#
filepaths_list["test_output_path"])
def all_sites_combo(raw_data_folder, species_list, outpath,

```

```

        data_outname="all_plots_combo.csv",
metrics_outname="all_plots_metrics.csv",
        top_species_outname="all_plots_top_species.csv",
top_val=10,
        metric_names=None, top_species_names=None,
species_outname="species_metrics_all_plots.csv"):
    if metric_names is None:
        metric_names = site_metrics
    if top_species_names is None:
        top_species_names = ["Top IVI Species. Native", "Top IVI Species.
All",
                            "Top C Species. Native"]

    process_start = datetime.datetime.now()
    print("Now preparing the combined data sheet and metrics.")

    filelist = get_raw_spreadsheets(raw_data_folder)
    species_sheet = pd.read_csv(species_list, header=0, index_col=0)

    big_quads_presence_index = []
    big_quads_coverage_index = []

    for file in filelist:

        filename = os.path.basename(file)
        site_num_whole = get_site_number(filename, tag1="site",
tag2="Site", delim="_")

        data_sheet = pd.read_excel(file, header=0, skiprows=[0],
index_col=0)
        data_sheet = data_sheet.fillna(0).astype(int)

        site_quadrat_index = []
        for quad in quadrat_number:
            pres = f"{quad}_presence_{site_num_whole}"
            site_quadrat_index.append(f"{quad}_presence_{site_num_whole}")
            big_quads_presence_index.append(pres)

            cov = f"{quad}_coverage_{site_num_whole}"
            site_quadrat_index.append(f"{quad}_coverage_{site_num_whole}")
            big_quads_coverage_index.append(cov)

        old_index = data_sheet.columns.values.tolist()
        zipped_index = list(zip(old_index, site_quadrat_index))

        new_quadrat_index = {}
        for old, new in zipped_index:
            new_quadrat_index[old] = new

        data_sheet = data_sheet.rename(columns=new_quadrat_index)
        species_sheet = pd.merge(species_sheet, data_sheet,
left_index=True, right_index=True, how="left")

        print("All sites successfully merged.")

        species_sheet = species_sheet.fillna(0)

        metrics_run = combined_metrics_run_single_sheet(species_sheet, top_val,
presence_iterator=big_quads_presence_index,

```

```

coverage_iterator=big_quads_coverage_index,

                                                                print_extra=True,
                                                                outpath=outpath, )

    metrics_total_step = metrics_run[0]
    top_species = metrics_run[1]

    big_metrics_total = pd.DataFrame([metrics_total_step],
columns=metric_names)

    big_metrics_top = pd.DataFrame([top_species],
columns=top_species_names)

    def indiv_species_metrics(species_sheet):

        species_list_index = species_sheet.index.values.tolist()
        species_metrics_columns = ["Frequency", "Coverage", ]
        species_metrics = pd.DataFrame(columns=species_metrics_columns,
index=species_list_index)

        for spec in species_list_index:
            spec_count = 0
            for plot in big_quads_presence_index:
                spec_count += species_sheet.loc[spec, plot]
            spec_row = species_sheet.loc[spec]
            mc_val = mean_cover_individual(spec_row,
coverage_iterator=big_quads_coverage_index,

midpoint_dict=coverage_midpoints)
            species_metrics.loc[spec] = {"Frequency": spec_count,
"Coverage": mc_val}

        return species_metrics

    species_sheet_native = species_sheet[species_sheet["Native Plant"] ==
1]

    species_metrics_all = indiv_species_metrics(species_sheet)
    species_outname_all = "all_" + species_outname
    species_metrics_native = indiv_species_metrics(species_sheet_native)
    species_outname_native = "native_" + species_outname

    for name_file, data in ((data_outname, species_sheet),
(metrics_outname, big_metrics_total),
(top_species_outname, big_metrics_top),
(species_outname_all, species_metrics_all),
(species_outname_native,
species_metrics_native)):
        outpath_file = outpath + name_file
        data.to_csv(outpath_file)

    generic_outpath = outpath + data_outname

    process_end = datetime.datetime.now()

    print(f"Metrics completed. Big data saved to:\n{generic_outpath}")
    print(f"Time elapsed: {str(process_end - process_start)}")

    return species_sheet, big_metrics_total, big_metrics_top,
species_metrics_all, species_metrics_native, generic_outpath

```

```

# This function will take a single zone data table and calculate metrics
based on a
# mixed model design; i.e., this will produce a table with a row of metrics
for every
# single plot.
# PLEASE NOTE THAT INTRODUCED SPECIES HERE IS CALCULATED WITH ONLY
IDENTIFIED SPECIES!
# zone_metric_test = mixed_model_one_zone(filepaths_list["test_raw_data"],
species_sheet, quadrat_presence,
#                                     coverage_midpoints,
outsave_yesno=True, outfolder=filepaths_list["output_path"])
# print(zone_metric_test)
def mixed_model_one_zone(zone_filepath, species_list_df, presence_iterator,
midpoint_dict, colnames,
outfolder=None, outname="mixed_model",
outsave_yesno=False,
native_col="Native Plant", cval_col="CoC"):

    filename = os.path.basename(zone_filepath)
    site_num_whole = get_site_number(filename)
    zoi_num = "zoi_" + site_num_whole[-2:]
    print(f"Now creating a mixed model spreadsheet for {zoi_num}...")

    zone_dataframe = pd.read_excel(zone_filepath, header=0, skiprows=[0],
index_col=0)
    zone_dataframe = zone_dataframe.fillna(0).astype(int)

    merged_zone_df_all_w_unident = pd.merge(zone_dataframe,
species_list_df, left_index=True, right_index=True, how="left")

    merged_zone_df_all_wo_unident = merged_zone_df_all_w_unident[
merged_zone_df_all_w_unident[native_col] != 3]
    merged_zone_df_native = merged_zone_df_all_w_unident[
merged_zone_df_all_w_unident[native_col] == 1]

    zone_metrics_df = pd.DataFrame(columns=colnames)

    for iterator in presence_iterator:
        roi_num = iterator[:3]
        cov_iterator = roi_num + "_coverage"
        index = f"{zoi_num}_{roi_num}"

        sp_rich_nat = richness_quads(merged_zone_df_native, iterator)
        sp_rich_all = richness_quads(merged_zone_df_all_w_unident,
iterator)

        perc_int = percent_introduced_quads(merged_zone_df_all_wo_unident,
iterator)

        c_mean_nat = mean_c_quads(merged_zone_df_native, iterator)
        c_mean_all = mean_c_quads(merged_zone_df_all_wo_unident, iterator)

        fqi_nat = c_mean_nat * math.sqrt(sp_rich_nat)
        fqi_all = c_mean_all * math.sqrt(sp_rich_nat)

        afqi = afqi_func(c_mean_nat, sp_rich_nat, sp_rich_all)

        encroach =
modified_plot_encroachment_metrics(merged_zone_df_all_w_unident,
cov_iterator, midpoint_dict)

```



```

        woody_encroach = encroach[1]

        values_list = [zoi_num, roi_num, sp_rich_nat, sp_rich_all,
perc_int, c_mean_nat,
                        c_mean_all, fqi_nat, fqi_all, afqi, woody_encroach]

        zone_metrics_df.loc[index,] = values_list

        if outsave_yesno is True:
            outpath = f"{outfolder}{outname}_{zoi_num}_plots.csv"
            zone_metrics_df.to_csv(outpath)

    print("Mixed model metrics calculated.\n")
    return zone_metrics_df

# This function takes mixed model plot metrics (from mixed_model_one_zone)
# and runs
# these metrics across every zone in a folder, and returns a single sheet
# with
# mixed model metrics for all thirty zones. Basically it runs metrics
# across
# every individual plot, and produces a consolidated data sheet.
# mixed_model_all_zones(filepaths_list["raw_data_path"],
# filepaths_list["species_list"], quadrat_presence,
# coverage_midpoints, filepaths_list["output_path"])
def mixed_model_all_zones(raw_data_folder, species_list_path,
presence_iterator, midpoint_dict,
                        outfolder,
outname="mixed_model_all_zones_metrics.csv",
                        outname_inner=None, outfolder_inner=None,
outsave_yesno_inner=False, native_col="Native Plant",
                        cval_col="CoC", colnames=None,):
    if colnames is None:
        colnames=["zoi", "plot", "sp_richness_native_mm",
"sp_richness_all_mm",
                    "percent_introduced_mm", "mean_c_native_mm",
"mean_c_all_mm",
                    "mean_based_fqi_native_mm", "mean_based_fqi_all_mm",
"mean_based_afqi_mm",
                    "woody_pest_encroachment_mm"]

    process_start = datetime.datetime.now()
    print("Now preparing the combined data sheet and metrics.\n")

    filelist = get_raw_spreadsheets(raw_data_folder)

    species_sheet = pd.read_csv(species_list_path, header=0, index_col=0)

    all_zones_df = pd.DataFrame(columns=colnames)

    for filepath in filelist:
        zone_df = mixed_model_one_zone(filepath, species_sheet,
presence_iterator, midpoint_dict,
                                colnames, outfolder=outfolder_inner,
outsave_yesno=outsave_yesno_inner,
                                outname=outname_inner,
native_col=native_col, cval_col=cval_col)
        all_zones_df = pd.concat([all_zones_df, zone_df])

    outpath = outfolder + outname

```

```

all_zones_df.to_csv(outpath)

process_end = datetime.datetime.now()
print(f"Mixed Model zone calculations complete.\nTotal elapsed time:
{str(process_end-process_start)}")

return all_zones_df

def manage_sum_func(row, col_tag, answer_dict, iter_list, ):
    manage_val = 0
    for typeof in iter_list:
        col_name = typeof + "_" + col_tag
        answer = row[col_name]
        answer_val = answer_dict[answer]
        manage_val += answer_val

    return manage_val

def modified_manage_sum_func(dataset, answer_dict_dict, outfile,
measure_tag=None, iter_list=None, results_cols=None,
zone_num=None, row_tag="zoi"):
    if measure_tag is None:
        measure_tag = ["frequency", "percent", "timeago"]
    if iter_list is None:
        iter_list = ["burn", "mowing", "seeding", "herbicide", "handweed"]
    if results_cols is None:
        results_cols = ["burn_sum", "mowing_sum", "seeding_sum",
"herbicide_sum", "handweed_sum"]
    if zone_num is None:
        zone_num = list(range(1, 15)) + list(range(17, 31))

    manage_sums_df = pd.DataFrame(columns=results_cols)

    for zone in zone_num:
        row_name = row_tag + str(zone)
        print(f"Running survey metrics for {row_name}")

        row = dataset.loc[row_name,]

        for typeof in iter_list:
            typeof_sum = 0
            for measure in measure_tag:
                col_name = typeof + "_" + measure
                answer = row[col_name]
                answer_val = answer_dict_dict[measure][answer]
                typeof_sum += answer_val

            results_name = typeof + "_sum"
            manage_sums_df.loc[row_name, results_name] = typeof_sum

    manage_sums_df.to_csv(outfile)

    return manage_sums_df

def staff_sum_func(row, answer_dict_list, iter_list, contractors_dict,
contractors_name="contractors_yesno",
col_tag="_number"):
    staff_val = 0

```

```

zipped_iterator = list(zip(iter_list, answer_dict_list))

for staff_type, staff_dict in zipped_iterator:
    col_name = staff_type + col_tag
    answer = row[col_name]
    if answer == "Don't know":
        return "err"
    answer_val = staff_dict[answer]
    staff_val += answer_val

contract = row[contractors_name]
staff_val += contractors_dict[contract]

return staff_val

def rescale(unscaled_num, new_min=1, new_max=10, old_min=0, old_max=80):
    perc = (unscaled_num - old_min) / (old_max - old_min)
    new_scale = ((new_max - new_min) * perc) + new_min
    return new_scale

# survey_responses_summary(filepaths_list["survey_responses"],
dict_of_keydicts, survey_zones_list,
#                             filepaths_list["test_output_path"])
def survey_responses_summary(survey_filepath, answers_dict, row_iterator,
outpath,
                                outname="survey_results_summary_metrics.csv",
                                row_tag="zoi_", result_column_names=None,
                                zone_area_col="zone_area",
site_area_col="site_area",
                                org_budget_modifier=0.25,
manage_techs_list=None,
                                staff_tags=None,
site_funds_col="funds_per_zone",
                                site_num_col="number_org_sites",
org_budget_col="org_operating_budget"):
    if result_column_names is None:
        result_column_names = ["freq_sum", "percent_sum", "timeago_sum",
"management_sum",
                                "manage_sum_per_hect_zone",
"manage_sum_per_hect_site",
                                "management_scale", "staff_sum",
"staff_scale",
                                "manage_w_staff_mult",
                                "manage_w_staff_mult_per_hect_zone",
                                "manage_w_staff_mult_per_hect_site",
                                "manage_staff_scale",
"site_budget_per_hect_zone", "site_budget_per_hect_site",
                                "org_budget_per_site",
"org_budget_per_hect_zone",
                                "org_budget_per_hect_site", ]
    if manage_techs_list is None:
        manage_techs_list = ["burn", "mowing", "seeding", "herbicide",
"handweed"]
    if staff_tags is None:
        staff_tags = ["permanent_staff", "seasonal_staff", "volunteers"]

process_start = datetime.datetime.now()
print("Summarizing survey metrics now.")

```

```

survey_responses_file = pd.read_excel(survey_filepath, skiprows=[0],
header=0, index_col=0)
survey_file_summary = pd.DataFrame(columns=result_column_names)

for zone in row_iterator:
    row_name = row_tag + str(zone)
    print(f"Running survey metrics for {row_name}")
    zone_area_val = survey_responses_file.loc[row_name, zone_area_col]
    site_area_val = survey_responses_file.loc[row_name, site_area_col]
    row = survey_responses_file.loc[row_name,]

    freq_sum = manage_sum_func(row, "frequency",
answers_dict["freq_dict"], manage_techs_list)
    percent_sum = manage_sum_func(row, "percent",
answers_dict["percent_dict"], manage_techs_list)
    timeago_sum = manage_sum_func(row, "timeago",
answers_dict["timeago_dict"], manage_techs_list)
    management_sum = freq_sum + percent_sum + timeago_sum
    management_sum_per_hect_zone = management_sum / zone_area_val
    management_sum_per_hect_site = management_sum / site_area_val
    management_scale = rescale(management_sum)

    staff_dict_list = [answers_dict["permstaff_dict"],
answers_dict["seasonstaff_dict"],
                        answers_dict["volunteer_dict"]]
    staff_sum = staff_sum_func(row, staff_dict_list, staff_tags,
answers_dict["contractor_dict"])

    if staff_sum != "err":
        staff_scale = rescale(staff_sum, old_max=26)
        manage_w_staff_mult = management_sum + (2 * staff_sum)
        manage_w_staff_mult_per_hect_zone = manage_w_staff_mult /
zone_area_val
        manage_w_staff_mult_per_hect_site = manage_w_staff_mult /
site_area_val
        manage_staff_scale = rescale(manage_w_staff_mult, old_max=132)
    else:
        staff_scale = "err"
        manage_w_staff_mult = "err"
        manage_w_staff_mult_per_hect_zone = "err"
        manage_w_staff_mult_per_hect_site = "err"
        manage_staff_scale = "err"

    if row[site_funds_col] != "Don't know":
        site_budget =
answers_dict["site_budget_dict"][row[site_funds_col]]
        site_budget_per_hect_zone = site_budget / zone_area_val
        site_budget_per_hect_site = site_budget / site_area_val
    else:
        site_budget_per_hect_zone = "err"
        site_budget_per_hect_site = "err"

    if row[site_num_col] != "Don't know" and row[org_budget_col] !=
"Don't know":
        site_num =
answers_dict["number_other_sites_dict"][row[site_num_col]]
        org_budget = org_budget_modifier *
answers_dict["org_budget_dict"][row[org_budget_col]]

        org_budget_per_site = org_budget / site_num

```



```

        org_budget_per_hect_zone = org_budget / zone_area_val
        org_budget_per_hect_site = org_budget / site_area_val
    elif row[site_num_col] == "Don't know" and row[org_budget_col] !=
"Don't know":
        org_budget = org_budget_modifier *
answers_dict["org_budget_dict"][row[org_budget_col]]

        org_budget_per_site = "err"
        org_budget_per_hect_zone = org_budget / zone_area_val
        org_budget_per_hect_site = org_budget / site_area_val
    else:
        org_budget_per_site = "err"
        org_budget_per_hect_zone = "err"
        org_budget_per_hect_site = "err"

    row_metrics_list = [freq_sum, percent_sum, timeago_sum,
management_sum, management_sum_per_hect_zone,
                        management_sum_per_hect_site,
                        management_scale, staff_sum, staff_scale,
manage_w_staff_mult,
                        manage_w_staff_mult_per_hect_zone,
manage_w_staff_mult_per_hect_site,
                        manage_staff_scale, site_budget_per_hect_zone,
site_budget_per_hect_site,
                        org_budget_per_site, org_budget_per_hect_zone,
org_budget_per_hect_site]

    survey_file_summary.loc[row_name,] = row_metrics_list

    print("All rows summarized.\n")
    out_filepath = outpath + outname
    survey_file_summary.to_csv(out_filepath)

    print(f"File saved to: {out_filepath}\n")
    process_end = datetime.datetime.now()
    print(f"Time elapsed: {str(process_end - process_start)}")

    return survey_file_summary

# ## BACKUPS AND CHANGELOG ## #
# v1.6 added mixed model organizational functions, which produce a single
# spreadsheet with summary metrics for each individual plot, to be used as
# a mixed-model alternative to normal statistics that use the average-plot
# values.
#
# v1.5 added functions to summarize the survey response information and
# generate summary metrics for those surveys.
# NOTE: the survey metrics functions are fairly hard-coded and would need
# to be
# revised to be useful in any other situation.
#
# v1.4 changed the way that gradient land use handles zeros and zero-
# values,
# so that ratios that would be infinity (i.e. divide by zero) instead
# replace with
# nan, which is later ignored by summary metrics. The alternative is to
# replace zeroes with 0.001 or similar, which is also a built-in function.
#
# v1.3 added gradient land use metrics and functions associated with that
# v1.1 Fixed handling of areas with 0 mean or median C values.

```

```
# v1.0 - The initial release version. Contains functions for binary land
use
# classification, all site metrics, and combining all the individual sites
# into a singular datasheet.
```

R-Based Statistical Analysis and Data Exploration

This includes the functions used to run statistical analyses and the tools I utilized to make statistical decisions regarding my analysis. It excludes the code for graph production and unstructured data exploration, especially data exploration that did not produce any viable information (e.g. Mixed Modeling).

```
Thesis_statistics_analysis.R
# Stats for Thesis Project, Andy Sima
# February-March 2025
# Version 1.6 Apr. 23, 2025 (Copied on June 2 2025)

# installing and loading packages-----
library(ggplot2)
library(car)
library(lmtest)
library(multcomp)
library(dplyr)
# install.packages("Rtools")
# library(Rtools)
# install.packages("ggpubr")
library(ggpubr)
library(grid)
library(gridExtra)
library(stringr)
library(gdata)
library(lme4)
# install.packages("lmerTest")
library(lmerTest)
# install.packages("sjPlot")
library(sjPlot)
library(tools)
library(effects)
#install.packages("pscl")
library(pscl)
#install.packages("tidyverse")
library(tidyverse)

# Establishing Global Variables-----

# setting working directory
# AT SCHOOL
mainDir <- "C:/Users/ansi0618/OneDrive/REAL_DOCUMENTS_LIVE/Stockholm
University Classwork/THESIS WORK/full_data_TOTAL/" # at school
# AT HOME
#mainDir <- "C:/Users/andre/OneDrive/REAL_DOCUMENTS_LIVE/Stockholm
University Classwork/THESIS WORK/full_data_TOTAL/" # at home
setwd(mainDir) #at school

prairie_df <- read.csv("dependent_variables_combined.csv") # Dependent
variables of the prairie plants
```

```

inde_df <- read.csv("independent_variables_combined.csv") # Independent
variables of land use, management, etc.
mixed_model_df <- read.csv("dependent_mixed_model_all_zones_metrics_v1-6_3-
14-25.csv") #mixed model summary metrics of all observations

inde_df[inde_df=="nan"] <- NA
inde_df[inde_df=="NaN"] <- NA
inde_df[inde_df=="err"] <- NA

combo_df_zone <- merge(prairie_df, inde_df, by="X")
combo_df_mixed <- merge(mixed_model_df, inde_df, by.x="zoi", by.y="X")

# All the Variables in Lists-----
mixed_model_list_w_zonenames <- c(
  "zoi",
  "plot",
  "sp_richness_native_mm",
  "sp_richness_all_mm",
  "percent_introduced_mm",
  "mean_c_native_mm",
  "mean_c_all_mm",
  "mean_based_fqi_native_mm",
  "mean_based_fqi_all_mm",
  "mean_based_afqi_mm",
  "woody_pest_encroachment_mm"
)

mixed_model_response_list_no_names <- c(
  "sp_richness_native_mm",
  "sp_richness_all_mm",
  "percent_introduced_mm",
  "mean_c_native_mm",
  "mean_c_all_mm",
  "mean_based_fqi_native_mm",
  "mean_based_fqi_all_mm",
  "mean_based_afqi_mm",
  "woody_pest_encroachment_mm"
)

response_variables_list <- c(
  "sp_richness_native_plots_avg",
  "sp_richness_all_plots_avg",
  "percent_introduced_sp_plots_avg",
  "mean_c_native_plots_avg",
  "mean_c_all_plots_avg",
  "mean_based_fqi_native_plots_avg",
  "mean_based_fqi_all_plots_avg",
  "mean_based_afqi_all_plots_avg",
  "shannon_diversity_index_native_site",
  "shannon_diversity_index_all_site",
  "simpson_diversity_index_native_site",
  "simpson_diversity_index_all_site"
)

final_response_variables_list <- c(
  "sp_richness_native_plots_avg",
  "percent_introduced_sp_plots_avg",
  "mean_based_fqi_native_plots_avg",
  "mean_based_afqi_all_plots_avg",
  "shannon_diversity_index_all_site"
)

```

```

final_response_Variables_list_mm <- c(
  "sp_richness_native_mm",
  "percent_introduced_mm",
  "mean_based_fqi_native_mm",
  "mean_based_afqi_mm"
)

response_variables_name_list <- c(
  "Native Species Richness, Plots Avg.",
  "All Species Richness, Plots Avg.",
  "% Introduced Sp., Plots Avg.",
  "Mean C Native Sp., Plots Avg.",
  "Mean C All Sp., Plots Avg.",
  "Native FQI, Plots Avg.",
  "All FQI, Plots Avg.",
  "AFQI, Plots Avg.",
  "Native Shannon Diversity",
  "All Shannon Diversity",
  "Native Gini-Simpson Diversity",
  "All Gini-Simpson Diversity"
)

final_response_variables_name_list <- c(
  "Native Species Richness, Plots Avg.",
  "% Introduced Sp., Plots Avg.",
  "Native FQI, Plots Avg.",
  "AFQI, Plots Avg.",
  "All Shannon Diversity"
)

site_response_variables_list <- c(
  "sp_richness_native_site",
  "sp_richness_all_site",
  "percent_introduced_sp_site",
  "mean_c_native_site",
  "mean_c_all_site",
  "mean_based_fqi_native_site",
  "mean_based_fqi_all_site",
  "mean_based_afqi_all_site"
)

all_categorical_predictor_list <- c(
  "land_class_binary_l3_site",
  "selfreport_num_methods_used",
  "selfreport_num_methods_used_grouped",
  "burn_yesno",
  "mowing_yesno",
  "seeding_yesno",
  "herbicide_yesno",
  "handweeding_yesno",
  "permanent_staff_number",
  "funds_per_zone",
  "funds_per_zone_grouped",
  "org_operating_budget",
  "org_operating_budget_grouped",
  "sector_type_pcsc",
  "special_legal_designations",
  "age_of_restoration_grouped",
  "remnant_restore_status_new"
)

```



```

valid_categorical_predictor_list <- c(
  "land_class_binary_l3_site",
  "selfreport_num_methods_used",
  "selfreport_num_methods_used_grouped",
  "mowing_yesno",
  "seeding_yesno",
  "handweeding_yesno",
  "permanent_staff_number",
  "funds_per_zone",
  "funds_per_zone_grouped",
  "org_operating_budget",
  "org_operating_budget_grouped",
  "sector_type_pcsc",
  "special_legal_designations",
  "age_of_restoration_grouped"
)

continuous_predictor_list <- c(
  "zoi_area_hectares_utm16n",
  "site_area_zoi_included_hectares",
  "dist_to_chicago_km",
  "grad_avg_percent_total_wild",
  "grad_slope_percent_total_wild",
  "grad_relative_slope_percent_total_wild",
  "grad_avg_percent_open_wild",
  "grad_slope_percent_open_wild",
  "grad_relative_slope_percent_open_wild",
  "grad_avg_percent_rural",
  "grad_slope_percent_rural",
  "grad_relative_slope_percent_rural",
  "grad_avg_percent_suburban",
  "grad_slope_percent_suburban",
  "grad_relative_slope_percent_suburban",
  "percent_total_wild_buff50m",
  "percent_open_wild_buff50m",
  "percent_rural_buff50m",
  "percent_suburban_buff50m",
  "ratio_rural_suburban_buff50m",
  "percent_total_wild_buff1500m",
  "percent_open_wild_buff1500m",
  "percent_rural_buff1500m",
  "percent_suburban_buff1500m",
  "ratio_rural_suburban_buff1500m",
  "percent_total_wild_buff3000m",
  "percent_open_wild_buff3000m",
  "percent_rural_buff3000m",
  "percent_suburban_buff3000m",
  "ratio_rural_suburban_buff3000m",
  "selfreport_manage_intensity",
  "burn_sum",
  "mowing_sum",
  "seeding_sum",
  "herbicide_sum",
  "handweed_sum",
  "staff_sum",
  "management_sum",
  "manage_w_staff_mult",
  "manage_w_staff_mult_per_hect_zone",
  "manage_w_staff_mult_per_hect_site",
  "org_budget_per_site",

```

```

"site_budget_per_hect_zone",
"site_budget_per_hect_site",
"restore_to_remnant_scale",
"age_of_restoration",
"age_of_restoration_w_remnant",
"bin2k_area_r_over_s_ratio_l3_site_log",
"grad_avg_ratio_totalwild_dominant_log",
"grad_avg_ratio_openwild_closedwild_log",
"grad_avg_ratio_rural_suburban_log",
"ratio_rural_suburban_buff50m_log",
"ratio_rural_suburban_buff1500m_log",
"ratio_rural_suburban_buff3000m_log"
)

important_predictor_variables_list <- c(
  "zoi_area_hecatres_utm16n",
  "grad_avg_percent_rural",
  "percent_rural_buff1500m",
  "grad_avg_ratio_rural_suburban_log",
  "restore_to_remnant_scale",
  "staff_sum",
  "management_sum",
  "handweed_sum",
  "site_budget_per_hect_zone"
)

# cleaning_data_func: Cleaning Data Function to be applied to each mixed
# model and zone-based dataset-----
# This only does anything to independent columns so it can be applied
# pretty easy
# to both mixed and zone
# It will introduce NAs by coercion but this is to be expected. Those are
# teh
# age of restorations where it says "remnant".
cleaning_data_func <- function(combo_df,
                               cont_pred_list=continuous_predictor_list,
                               cat_pred_list=all_categorical_predictor_list,
                               resp_vari_list=response_variables_list){

  print("Now cleaning file...")

  combo_df[combo_df=="Don't know"] <- NA

  # Log Transforming Ratio Columns to be Better About Magnitude-----
  ---
  log_columns <- c("bin2k_area_r_over_s_ratio_l3_site",
                  "grad_avg_ratio_totalwild_dominant",
                  "grad_avg_ratio_openwild_closedwild",
                  "grad_avg_ratio_rural_suburban",
                  "ratio_rural_suburban_buff50m",
                  "ratio_rural_suburban_buff1500m",
                  "ratio_rural_suburban_buff3000m")

  log_list <- select(combo_df, "X", all_of(log_columns))
  log_list <- log_list %>%
    mutate(across(log_columns, as.numeric))

  new_logs <- log_list %>%
    mutate(across(log_columns, log))

```

```

new_col_names <- list("X")
for (col in log_columns){
  new_col <- list(paste(col, "_log", sep=""))
  new_col_names <- append(new_col_names, new_col)
}

names(new_logs) <- new_col_names

new_logs[new_logs==--Inf] <- NA

combo_df <- merge(combo_df, new_logs, by="X")

# Setting certain columns to factors-----
factor_list <- cat_pred_list

combo_df <- combo_df %>% mutate(selfreport_num_methods_used_grouped =
case_when(
  selfreport_num_methods_used %in% c(1, 2, 3) ~ "3 or Less",
  selfreport_num_methods_used %in% c(4) ~ "4 Methods",
  selfreport_num_methods_used %in% c(5, 6) ~ "5 to 6",
  TRUE ~ NA
))

combo_df <- combo_df %>% mutate(funds_per_zone_grouped = case_when(
  funds_per_zone %in% c("Less than $100,000") ~ "Less than $100,000",
  funds_per_zone %in% c("Between $500,000 and $1 million",
    "Between $100,000 and $500,000") ~ "Greater than
$100,000",
  TRUE ~ NA
))

combo_df <- combo_df %>% mutate(org_operating_budget_grouped = case_when(
  org_operating_budget %in% c("between $150 and $200 million", "Between
$600 million and $1 billion",
    "greater than $1 billion") ~ "Greater than
$150 Million",
  org_operating_budget %in% c("Between $100,000 and $500,000", "Between
$1 and $2.5 Million",
    "between $2.5 and $5 million") ~ "Less than
$5 Million",
  org_operating_budget %in% c("between $15 million and $20 million",
    "between $20 and $30 million",
    "between $50 and $75 million", "between
$100 and $150 million") ~ "Between $5 and $150 Million",
  TRUE ~ NA
))

combo_df <- combo_df %>% mutate(age_of_restoration_grouped = case_when(
  age_of_restoration_w_remnant > 0 & age_of_restoration_w_remnant < 31 ~
"Younger than 30 yrs",
  age_of_restoration_w_remnant > 31 & age_of_restoration_w_remnant < 100
~ "Between 30 and 100 yrs",
  age_of_restoration_w_remnant %in% c(187) ~ "Remnant",
  TRUE ~ NA
))

for (factor in factor_list){
  combo_df[[factor]] <- as.factor(combo_df[[factor]])
}

```

```

    combo_df$last_date_sampled_asdate <- as.Date(combo_df$last_date_sampled,
format = "%d/%m/%Y")
    combo_df$last_date_sampled_asnumeric <-
as.numeric(combo_df$last_date_sampled_asdate)

#Making sure everything is numeric that should be-----
suppressWarnings({
  for (colname in cont_pred_list){
    combo_df[[colname]] <- as.numeric(combo_df[[colname]])
  }
  # Introducing NAs by coercion is for age_of_restoration without
remnants, it turns
  # the "remnant" string into an NA
  #Suppress Warnings has been introduced to stop this error from coming
up.
})

for (colname in resp_vari_list){
  combo_df[[colname]] <- as.numeric(combo_df[[colname]])
}

print("File cleaned.")
return(combo_df)
}

# Cleaning all the Datasets-----
combo_df_zone <- cleaning_data_func(combo_df_zone)
combo_df_mixed <- cleaning_data_func(combo_df_mixed,
resp_vari_list=mixed_model_response_list_no_names)

#Removing Trelease and Midewin Prairie AND ALSO GROUPING FACTORS-----
--
combo_df_zone_droptrelease <- combo_df_zone[combo_df_zone$zoi_number != 3,]
combo_df_mixed_droptrelease <- combo_df_mixed[combo_df_mixed$zoi !=
"zoi_03",]

# Removing Midewin is important for some calculations between it is empty
of all management
# data. Hopefully you don't have to use it much.
combo_df_zone_dropmidewin <- combo_df_zone %>% filter(!(zoi_number %in%
c(15, 16)))
combo_df_mixed_dropmidewin <- combo_df_mixed %>% filter(!(zoi %in%
c("zoi_15", "zoi_16")))

#GROUPING FACTORS
combo_df_zone_onlyrural_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$land_class_binary_13_
site == "Rural",]
combo_df_zone_onlysuburban_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$land_class_binary_13_
site == "Suburban",]

combo_df_zone_handweedyes_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$handweeding_yesno ==
"yes",]

```



```

combo_df_zone_handweedno_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$handweeding_yn ==
"no",]

combo_df_zone_public_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$sector_type_pc ==
"Public",]
combo_df_zone_private_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$sector_type_pc ==
"Private",]

combo_df_zone_yeslegal_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$special_legal_designa
tions == "Yes",]
combo_df_zone_nolegal_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$special_legal_designa
tions == "No",]

combo_df_zone_orgbudget_5mil_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$org_operating_budget_
grouped == "Less than $5 Million",]
combo_df_zone_orgbudget_5150mil_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$org_operating_budget_
grouped == "Between $5 and $150 Million",]
combo_df_zone_orgbudget_greater150_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$org_operating_budget_
grouped == "Greater than $150 Million",]

combo_df_zone_agerest_remnant_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$age_of_restoration_gr
ouped == "Remnant",]
combo_df_zone_agerest_30yrs_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$age_of_restoration_gr
ouped == "Younger than 30 yrs",]
combo_df_zone_agerest_30100yrs_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$age_of_restoration_gr
ouped == "Between 30 and 100 yrs",]

combo_df_zone_nummethod_less3_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$selfreport_num_method
s_used_grouped == "3 or Less",]
combo_df_zone_nummethod_4meth_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$selfreport_num_method
s_used_grouped == "4 Methods",]
combo_df_zone_nummethod_5or6_dropt <-
combo_df_zone_droptrelease[combo_df_zone_droptrelease$selfreport_num_method
s_used_grouped == "5 to 6",]

print("All Factors now grouped and cleaned.")

#Functions that are Useful-----
#Empty section here

# count_factor: A function to count the number of values in a factor.-----
-----
count_factor <- function(dataset, column){
  group_col <- sym(column)
  dataset %>%
    group_by(!group_col) %>%
    summarise(no_rows = n(), .groups="drop")

```

```

}

#count_factor(combo_df_zone_droptrelease, "org_operating_budget_grouped")

# simple_normal_test: A Function to determine if a dataset is normal or
not, but simplified-----
# returns TRUE if Normal, FALSE if not normal (pval less than 0.05)
# Also returns the pvalue as the second term.
simple_normal_test <- function(data, column){
  norm <- shapiro.test(data[[column]])
  pval_first <- norm$p.value

  yesno <- ifelse(pval_first>0.05, TRUE, FALSE)
  pval <- format(round(pval_first, 5), nsmall=5)

  return(c(yesno, pval))
}

# advanced_normal_test: A Function and Codeblock to get normality and
Pvalues from all data in a dataset.-----
# This one is advanced and returns a string answer that says either yes or
no plus
# the pvalue of the normal test.
advanced_normal_test <- function(data, column){
  norm <- shapiro.test(data[[column]])
  pval_first <- norm$p.value

  yesno <- ifelse(pval_first>0.05, "Yes", "No")
  pval <- format(round(pval_first, 5), nsmall=5)

  print(paste(column, ": ", yesno, " pval=", pval, sep=""))

  answer <- paste(yesno, " pval=", pval, sep="")

  qqPlot(data[[column]])
  return(answer)
}

# get_normal_table: A Function to get a table out normality and print it---
-----
get_normal_table <- function(zone_df, mixedm_df, resp_vari_list,
pred_cont_list, mixed_model_vari_list,
                           printfile=FALSE,
outname="normality_test_w_mm.csv"){
  normal_dataframe <- data.frame(pval="txt")

  for (colname in resp_vari_list){
    answ <- advanced_normal_test(zone_df, colname)
    normal_dataframe[colname,] <- answ
  }

  for (colname in pred_cont_list){
    answ <- advanced_normal_test(zone_df, colname)
    normal_dataframe[colname,] <- answ
  }

  for (colname in mixed_model_vari_list){
    answ <- advanced_normal_test(mixedm_df, colname)

```

```

    normal_dataframe[colname,] <- answ
  }

  if (printfile == TRUE){
    #WROTE THIS ALREADY SO WE DON'T NEED TO AGAIN
    write.csv(normal_dataframe, outname)
  }

  return(normal_dataframe)
}

# normal_df <- get_normal_table(combo_df_zone, combo_df_mixed,
response_variables_list,
#                               continuous_predictor_list,
mixed_model_list_no_names,
#                               printfile = TRUE)

# conti_conti_testing: A Function to get the GLM data out of a combination
of continuous variables-----
# This can be altered for normal or non-normal RESPONSE data by changing
the method.
# "pearson" is Normal distribution. (response data)
# "spearman" is Not Normal distribution. (response data)
conti_conti_testing <- function(response, predictor, data, normality=NULL){
  conti_dummy_model <- glm(data[[response]]~data[[predictor]], data=data)

  if (is.null(normality)){
    normality <- simple_normal_test(data, response)[1]
  }

  glm_dummy = Anova(conti_dummy_model)
  glm_chi = glm_dummy$`LR Chisq`
  glm_df = glm_dummy$Df
  glm_p = glm_dummy$`Pr(>Chisq)`

  method <- ifelse(normality, "pearson", "spearman")

  suppressWarnings({
    cor_test <- cor.test(x=data[[response]], y=data[[predictor]], method =
method)
    cor_p = cor_test$p.value
    cor_rho = cor_test$estimate
  })

  return(c(glm_chi, glm_df, glm_p, method, cor_p, cor_rho))
}

# conti_conti_testing_LM: A function to compare variables but uses lm
instead-----
# A modified version of the function above that returns variables from an
LM model
# Uses summary(lm())
# instead of anova(glm())
conti_conti_testing_LM <- function(response, predictor, data,
normality=NULL){
  conti_dummy_model <- lm(data[[response]]~data[[predictor]], data=data)

  if (is.null(normality)){
    normality <- simple_normal_test(data, response)[1]
  }

```

```

}

anova_dummy <- Anova(conti_dummy_model)
lm_f <- anova_dummy$`F value`[1]
lm_df <- paste("num:", anova_dummy$Df[1], ", den:", anova_dummy$Df[2])
lm_p <- anova_dummy$`Pr(>F)`[1]
lm_rsqr <- summary(conti_dummy_model)$r.squared

method <- ifelse(normality, "pearson", "spearman")

suppressWarnings({
  cor_test <- cor.test(x=data[[response]], y=data[[predictor]], method =
method)
  cor_p = cor_test$p.value
  cor_rho = cor_test$estimate
})

return(c(lm_f, lm_df, lm_p, lm_rsqr, method, cor_p, cor_rho))
}

#conti_conti_testing_LM("sp_richness_native_plots_avg",
"ratio_rural_suburban_buff50m_log", combo_df_zone_droptrelease, TRUE)

# conti_categori_testing: A Function to get the relevant data out of a
continuous to Categorical variable.-----
# DOES NOT do a Tukey test on the data. Only shows the overall p value,
have to dig into
# the specifics later.
#PUT RESPONSE VARIABLE FIRST.
conti_categori_testing <- function(response, predict, data, normality=NULL,
glm_ynsno=TRUE){

  if (is.null(normality)){
    normal <- simple_normal_test(data, response)[1]
  } else {
    normal <- normality
  }

  length <- length(levels(data[[predict]]))

  if(normal==TRUE & length==2){
    test_type <- "t.test"
    testes <- t.test(data[[response]]~data[[predict]])
    test_val <- testes$statistic
    test_df <- testes$parameter
    test_pval <- testes$p.value
  }else{
    if(normal==FALSE & length==2){
      test_type <- "wilcox"
      suppressWarnings({
        testes <- wilcox.test(data[[response]]~data[[predict]])
        test_val <- testes$statistic
        test_df <- NA
        test_pval <- testes$p.value
      })
    }else{
      if(normal==TRUE & length>2){
        test_type <- "Anova"
        aov_dummy_model <- Anova(aov(data[[response]]~data[[predict]]))
        test_val <- aov_dummy_model$`F value`[1]
      }
    }
  }
}

```



```

    test_df <- aov_dummy_model$Df[1]
    test_pval <- aov_dummy_model$`Pr(>F)`[1]
  }else{
    if(normal==FALSE & length>2){
      test_type <- "Kruskal"
      testes <- kruskal.test(data[[response]]~data[[predict]])
      test_val <- testes$statistic
      test_df <- testes$parameter
      test_pval <- testes$p.value
    }
  }
}

if (glm_yesno){
  return(c(test_val, test_df, test_pval, test_type, NA, NA))
} else {return(c(test_val, test_df, test_pval, NA, test_type, NA, NA))}
}

# all_data_overview: A function that will compare the p-values of different
combinations of continuous data-----
#This function assumes that all datasets fed into it are NORMAL (or at
least close enough
# to be handled by Anova and GLM) and that they are CONTINUOUS DATA (but
does return values of non-normal things too)
#This function will distinguish between categorical predictors and
continuous predictors,
# But still only uses ANOVA and GLM. Remember to look at the interesting
ones more thoroughly!
# IT WILL RETURN A COR.TEST VALUE FOR NORMAL OR NON-NORMAL DATA.
# This Function produces TWO separate spreadsheets. They are:
# 1. A large table comparing all possible combinations of (relevant)
variables, but ONLY p-values.
# 2. A spreadsheet of a more intricate breakdown of every combination of
data.
# This spreadsheet contains columns containing the following values:
# Response, predictor, categorical or continuous, normality,
# test_Val, test_df, test_pval, test_type or method, and cor.pval an
cor.rho for continuous data.
#RESPONSE VARIABLE GOES FIRST!
#ROW GOES FIRST! when calling dataframes
# This function has been modified to also accept the LM method for testing.
#PLEASE NOTE: the overview file does not distinguish between normal or not
normal. So it may not accurately
# depict trends or significances if the data is not normal.
# I FIXED THIS IN GENERIC GROUPING. May be fixed here too?
#We are serving Cont here.
col_names <- c("response","predictor","cati_or_conti_pred", "normality",
"test_val", "DF", "p_val", "method", "cor_pval", "cor_rho")
col_names_lm <- c("response","predictor","cati_or_conti_pred", "normality",
"test_val", "DF", "p_val", "r_square", "method", "cor_pval", "cor_rho")

all_data_overview <- function(dataset, response_list, cat_pred_list,
cont_pred_list, colum_names,
                                outsave=TRUE,
outname="all_variable_combinations_zones", pval_place=3,
                                method="glm"){
  df_view_row <- response_list
  df_view_col <- c(cat_pred_list, cont_pred_list)

```

```

df_view_matrix <- matrix(ncol=length(df_view_col),
nrow=length(df_view_row))
df_view <- data.frame(df_view_matrix)
colnames(df_view) <- df_view_col
rownames(df_view) <- df_view_row

df_all_matrix <- matrix(ncol=length(column_names),nrow=0)
df_all <- data.frame(df_all_matrix)
colnames(df_all) <- column_names

print("Beginning combination of all variables...")
cat("\n")

for (resp_vari in response_list){
  print("Now looking at Response Variable:")
  print(resp_vari)

  normality <- simple_normal_test(dataset, resp_vari)[1]

  if (length(cat_pred_list) != 0){
    for (pred_cat_vari in cat_pred_list){
      first_list <- c(resp_vari, pred_cat_vari, "categorical", normality)

      if (method == "glm"){
        categorical <- conti_categorical_testing(resp_vari, pred_cat_vari,
dataset, normality)
      } else {
        if (method == "lm"){
          categorical <- conti_categorical_testing(resp_vari, pred_cat_vari,
dataset, normality, glm_yesno = FALSE)
        } else {stop("Not a valid method")}
      }

      full_list <- c(first_list, categorical)

      thisrow <- paste(resp_vari, "_and_", pred_cat_vari, sep="")

      df_all[thisrow, ] <- full_list

      df_view[resp_vari, pred_cat_vari] <- categorical[pval_place]
    }
  }

  if (length(cont_pred_list) != 0){
    for (pred_cont_vari in cont_pred_list){
      first_list <- c(resp_vari, pred_cont_vari, "continuous", normality)

      if (method == "glm"){
        continuous <- conti_conti_testing(resp_vari, pred_cont_vari,
dataset, normality)
      } else {
        if (method == "lm"){
          continuous <- conti_conti_testing_LM(resp_vari, pred_cont_vari,
dataset, normality)
        } else {stop("Not a valid method")}
      }

      full_list <- c(first_list, continuous)

      thisrow <- paste(resp_vari, "_and_", pred_cont_vari, sep="")

```

```

    df_all[thisrow, ] <- full_list

    df_view[resp_vari, pred_cont_vari] <- continuous[pval_place]

  }
}

print("Response variable finished.")
cat("\n")
}

print("All response variables analyzed against all predictor variables.")
num_tests <- length(response_list) * length(c(cat_pred_list,
cont_pred_list))
print(paste("There were a total of:", num_tests, "combinations."))

if (outsave==TRUE){
  outname_dfview <- paste(outname, "_overview.csv", sep="")
  outname_dfall <- paste(outname, "_all_metrics.csv", sep="")

  write.csv(df_view, outname_dfview)
  write.csv(df_all, outname_dfall)

  print("File name is:")
  print(outname_dfall)
}

print("Combinations complete. I love you.")

return(c(df_view, df_all))
}

# EXAMPLES: Using the Above Functions to Do Something
# all_data_zone_dfs <- all_data_overview(combo_df_zone,
response_variables_list, valid_categorical_predictor_list,
continuous_predictor_list,
#                                     col_names, outsave=FALSE)
#
# all_data_mm_dfs <- all_data_overview(combo_df_mixed,
mixed_model_response_list_no_names, valid_categorical_predictor_list,
continuous_predictor_list,
#                                     col_names,
outname="all_variable_combinations_mixedmodel", outsave=FALSE)
#
# all_data_zone_dfs_droptrelease <-
all_data_overview(combo_df_zone_droptrelease, response_variables_list,
valid_categorical_predictor_list, continuous_predictor_list,
#                                     col_names,
outname="all_variable_combinations_zones_droptrelease", outsave=FALSE)
#
# all_data_mm_dfs_droptrelease <-
all_data_overview(combo_df_mixed_droptrelease,
mixed_model_response_list_no_names, valid_categorical_predictor_list,
continuous_predictor_list,
#                                     col_names,
outname="all_variable_combinations_mixedmodel_droptrelease", outsave=FALSE)
#

```

```

# all_data_site_dfs_droptrelease <-
all_data_overview(combo_df_zone_droptrelease, site_response_variables_list,
valid_categorical_predictor_list,
#
continuous_predictor_list, col_names, outsave=FALSE,
outname="all_variable_combinations_SITEWIDE")
#
# all_data_zone_df_droptrelease_LMLOOK <-
all_data_overview(combo_df_zone_droptrelease, response_variables_list,
valid_categorical_predictor_list, continuous_predictor_list,
#
col_names_lm,
outsave = TRUE,
outname="LMNOW_all_variable_combinations_zones_droptrelease_LMversion",
#
method="lm")

# generic_glm_graph: Functions to Quickly and Easily create GLM graphs for
demonstration of data comparisons.-----
#Also produces the relevant statistics for the combination.
#This one optionally lets you group by some subset as well, though this may
not be the most
#Statistically meaningful.
generic_glm_graph <- function(dataset, response, predictor, resp_name,
pred_name, fill_col=NULL,
group_name="Land Class Type", seBars=TRUE,
smooth_method="glm", group_colors=c("blue", "red"),
save_yesno=FALSE,
save_name="generic_glm_graph.png", note="note",
subfolder="results_statistics/graphs/",
three_group_colors=c("blue", "red", "black"),
long_note=NULL){
  new_plot <- ggplot(data=dataset, aes(y=.data[[response]],
x=.data[[predictor]])) +
    geom_point() +
    geom_smooth(method=smooth_method, se=seBars,) +
    labs(x=pred_name, y=resp_name) +
    theme_classic() +
    theme(axis.text = element_text(size = 12),
          axis.title.x = element_text(size = 14),
          axis.title.y = element_text(size = 14),
          plot.margin = margin(0.5, 0.5, 0.5, 0.5, unit="cm"))

  if (!is.null(fill_col)) {
    if (length(group_colors) == length(levels(dataset[[fill_col]]))){
      new_plot <- new_plot + aes(color=.data[[fill_col]]) +
        scale_color_manual(group_name, values=group_colors)
    } else {
      new_plot <- new_plot + aes(color=.data[[fill_col]]) +
        scale_color_manual(group_name, values=three_group_colors)
    }
  }

  normality <- simple_normal_test(dataset, response)[1]

  cont_test <- conti_conti_testing(response, predictor, dataset, normality)
  if (normality == TRUE){
    return_list <- paste("Resp: ", response, " Pred: ", predictor,
                        "\n", note, " smoothing: ", smooth_method, "
stats for un-grouped, normal,"

```



```

        "\n", "Anova GLM, chi:", cont_test[1], ", df:",
cont_test[2], ", p:", cont_test[3])
    } else {
        return_list <- paste("Resp: ", response, " Pred: ", predictor,
        "\n", note, ", smoothing: ", smooth_method, ",
stats for un-grouped, not normal,",
        "\n", "Spearman Cor.test, rho:", cont_test[6], ",
p:", cont_test[5])
    }

    if (!is.null(long_note)){
        return_list <- paste(return_list, "\n", "GROUPING: ", long_note)
    }

    new_plot <- new_plot + labs(caption=return_list)

    if (save_yesno == TRUE){
        savename <- paste(subfolder, save_name, sep="")
        suppressWarnings({ggsave(savename, new_plot, width=9, height=6,
units="in")})
        print("Graph successfully saved.")
    }

    return(new_plot)
}
# example_plot <- generic_glm_graph(combo_df_zone_droptrelease,
"sp_richness_native_plots_avg", "grad_avg_ratio_rural_suburban_log",
# "Native Species Richness, Plots
Avg.", "Average Ratio of Rural to Suburban (log)",
fill_col="land_class_binary_l3_site",
# group_name="Land Class Type")

# generic_boxplot_graph: Functions to create quick and easy BOXPLOT graphs
for demonstrations of data comparisons-----
#Also produces the relevant statistics for the combination.
#Does NOT do a tukey test for p values of multiple groups
#Format of the scale_labels variable must be like this:
#c("Rural"="Rural (n=14)", "Suburban"="Suburban (n=15)")
generic_boxplot_graph <- function(dataset, response, predictor, resp_name,
pred_name,
                                scale_labels=NULL, save_yesno=FALSE,
save_name="generic_boxplot_graph.png",
                                note="note",
subfolder="results_statistics/graphs/"){
    new_plot <- ggplot(data=dataset, aes(x=.data[[predictor]],
y=.data[[response]])) +
        geom_boxplot(fill="lightgrey") +
        labs(x=pred_name, y=resp_name) +
        theme_classic() +
        theme(axis.text = element_text(size = 12),
axis.title.x = element_text(size = 14),
axis.title.y = element_text(size = 14),
plot.margin = margin(0.5, 0.5, 0.5, 0.5, unit="cm"))

    if (!is.null(scale_labels)){
        new_plot <- new_plot + scale_x_discrete(labels=scale_labels)
    }

    normality = simple_normal_test(dataset, response)[1]

```

```

cat_test = conti_categori_testing(response, predictor, dataset,
normality)
return_list = paste("Resp: ", response, " Pred: ", predictor,
                    "\n", note, " normality: ", normality, " ",
                    "\n", cat_test[4], " val: ", cat_test[1], " df: ",
cat_test[2], " p: ", cat_test[3], sep="")

new_plot <- new_plot + labs(caption=return_list)

if (save_yesno == TRUE){
  savename <- paste(subfolder, save_name, sep="")
  suppressWarnings({ggsave(savename, new_plot, width=9, height=6,
units="in")})
  print("Graph successfully saved.")
}

return(new_plot)
}
# example_plot2 <- generic_boxplot_graph(combo_df_zone,
"sp_richness_native_plots_avg", "land_class_binary_l3_site",
# "Native Species Richness, Plots
avg.", "Land Class Type", save_yesno = TRUE)

# all_response_graphs: A Function to run all the response variables against
ONE of the predictor variables-----
# And optionally produce graphs for all of them.
all_response_graphs <- function(dataset, resp_vari_list,
resp_vari_name_list, pred_vari, pred_vari_name,
categorical_tf, subfolder_full,
outsave_yesno=TRUE,
note="Drop Trelease",
outlier_note="dropTr", smoothmethod="glm",
print_yesno=FALSE){

  final_list <- c()

  for (resp_vari in resp_vari_list){
    name_index <- which(resp_vari_list == resp_vari)
    resp_name <- resp_vari_name_list[name_index]
    savename = paste(resp_vari, "_x_", pred_vari, "_", outlier_note,
".png", sep="")

    if (categorical_tf == TRUE){
      dummy_plot <- generic_boxplot_graph(dataset, resp_vari, pred_vari,
resp_name,
pred_vari_name,
save_yesno=outsave_yesno,
save_name=savename,
subfolder=subfolder_full, note=note)
      if (print_yesno){suppressWarnings({print(dummy_plot)})}

    } else {
      dummy_plot <- generic_glm_graph(dataset, resp_vari, pred_vari,
resp_name,
pred_vari_name,
save_yesno=outsave_yesno, smooth_method = smoothmethod,
save_name=savename,
subfolder=subfolder_full, note=note)
      if (print_yesno){suppressWarnings({print(dummy_plot)})}
    }
  }
}

```

```

    final_list[[name_index]] <- paste(subfolder_full, savename, sep="")
  }

  print("All Graphs produced for predictor variable.")
  if (print_yesno) {print(final_list)}

  return(final_list)
}

# example <- all_response_graphs(combo_df_zone_droprelease,
response_variables_list, response_variables_name_list,
#                               "ratio_rural_suburban_buff1500m_log",
"Ratio Rural-Suburban at 1500m (Log)",
#                               "results_statistics/graphs/handweed/",
categorical_tf = FALSE,
#                               outsave_yesno=TRUE,
#                               print_yesno = FALSE)

# all_resp_and_pred_graphs: A Function that takes the above functions and
# produces graphs for EVERY SINGLE-----
# POSSIBLE COMBINATION!
# Utter madness. My hubris will be my downfall.
# This returns a list of filepaths to the created graphs.
all_resp_and_pred_graps <- function(dataset, resp_vari_list,
resp_vari_name_list, cat_pred_list,
                                cont_pred_list,

subfolder_start="results_statistics/graphs/", main_dir=mainDir,
                                outsave_yesno=TRUE, note="Drop
Trelease", outlier_note="dropTr",
                                smoothmethod="glm", print_yesno=FALSE,
folder_tag=""){
  filelist <- c()

  make_file <- function(folder_vari_name){
    subfolder_full <- paste(subfolder_start, folder_vari_name, folder_tag,
"/", sep="")
    create_folder <- file.path(main_dir, subfolder_full, sep="")
    if (!dir.exists(create_folder)) {dir.create(create_folder)}
    return(subfolder_full)
  }

  if (length(cat_pred_list) != 0) { for (cat_pred in cat_pred_list){
    subfolder_inner <- make_file(cat_pred)
    print(paste("Now making graphs for ", cat_pred, sep=""))
    new_name <- toTitleCase(gsub("_", " ", cat_pred))
    dummy_list <- all_response_graphs(dataset, resp_vari_list,
resp_vari_name_list, cat_pred,
                                new_name, subfolder_inner,
categorical_tf=TRUE,
                                note = note, outlier_note =
outlier_note,
                                print_yesno = print_yesno,
outsave_yesno = outsave_yesno)
    filelist <- c(filelist, dummy_list)
  }}

  if (length(cont_pred_list) != 0) { for (cont_pred in cont_pred_list){

```

```

    subfolder_inner <- make_file(cont_pred)
    print(paste("Now making graphs for ", cont_pred, sep=""))
    new_name <- toTitleCase(gsub("_", " ", cont_pred))
    dummy_list <- all_response_graphs(dataset, resp_vari_list,
resp_vari_name_list, cont_pred,
                                new_name, subfolder_inner,
categorical_tf=FALSE,
                                note = note, outlier_note =
outlier_note,
                                print_yesno = print_yesno,
outsave_yesno = outsave_yesno,
                                smoothmethod = smoothmethod)
    filelist <- c(filelist, dummy_list)
  }}

count <- length(filelist)
cat("\n")
print(paste(count, "Graphs created. "))
print("Done. All graphs created. I love you.")
return(filelist)
}

#EXAMPLES!
# drop_trelease_allgraph <-
all_resp_and_pred_graps(combo_df_zone_droptrelease,
response_variables_list, response_variables_name_list,
#
valid_categorical_predictor_list, continuous_predictor_list,
outsave_yesno=TRUE,
#
                                print_yesno=FALSE)
#
# keep_trelease_allgraph <- all_resp_and_pred_graps(combo_df_zone,
response_variables_list, response_variables_name_list,
#
valid_categorical_predictor_list, continuous_predictor_list,
outsave_yesno=TRUE,
#
                                print_yesno=FALSE, note
= "KEEP Trelease", outlier_note = "keepTr")
#
# just_restoration_cat <- c("age_of_restoration_grouped")
# drop_trelease_restorecat <-
all_resp_and_pred_graps(combo_df_zone_droptrelease,
response_variables_list, response_variables_name_list,
#
                                just_restoration_cat,
c(), outsave_yesno=TRUE,
#
                                print_yesno=FALSE)
#
# keep_trelease_restorecat <- all_resp_and_pred_graps(combo_df_zone,
response_variables_list, response_variables_name_list,
#
                                just_restoration_cat, c(), outsave_yesno=TRUE,
#
                                print_yesno=FALSE, note = "KEEP Trelease",
outlier_note = "keepTr")

# generic_grouping_glm_plots_ALL: FUNCTION to stupidly and simply get
variables from grouping all of the different-----
# metrics together.
# THIS IS A REALLY SIMPLE STUPID FAST TOOL, not to be used heavily.
# ACTUALLY I UPDATED IT SO THAT IT WORKS BETTER
# but not too mcuh better: it only returns the final graph that is made.

```



```

generic_grouping_glm_plots_ALL <- function(predictor, pred_name,
first_stats_group = combo_df_zone_onlyrural_dropt, second_stats_group =
combo_df_zone_onlysuburban_dropt,
                                first_stats_name =
"onlyrural_dropt", second_stats_name = "onlysuburban_dropt", fulldataset =
combo_df_zone_droptrelease,
                                third_stats_group = NULL,
third_stats_name = NULL, short_note="DropTr,",
group_col="land_class_binary_l3_site", save_yn=FALSE, print_yn=FALSE,
                                folder_tag=NULL,
main_dir=mainDir){
  if (!is.null(folder_tag)){
    subfolder_full <- paste("zzz_GROUPED_", folder_tag, "/", sep="")
    create_folder <- file.path(main_dir, "results_statistics/graphs",
subfolder_full, sep="")
    if (!dir.exists(create_folder)) {dir.create(create_folder)}
  }

  for (resp_vari in final_response_variables_list){
    resp_file = substr(resp_vari, start=1, stop=16)
    pred_file = substr(predictor, start=1, stop=16)
    group_col_file = substr(group_col, start=1, stop=16)

    if (!is.null(folder_tag)){
      save_name = paste(subfolder_full, resp_file, "_x_", pred_file,
"_GROUP_", group_col_file, ".png", sep="")

    } else{
      save_name = paste(resp_file, "_x_", pred_file, "_GROUP_",
group_col_file, ".png", sep="")
    }

    get_stats <- function(stats_group, stats_name){
      stats_normal <- simple_normal_test(stats_group, resp_vari)[1]
      stats_dummy <- conti_conti_testing(resp_vari, predictor, stats_group,
stats_normal)

      if (stats_normal){
        printer <- paste(stats_name, "Normal:", stats_dummy[1],
stats_dummy[2], stats_dummy[3], sep=", ")
      } else {
        printer <- paste(stats_name, "NOT Normal:", stats_dummy[6],
stats_dummy[5], sep=", ")
      }

      return(printer)
    }

    printer_1 <- get_stats(first_stats_group, first_stats_name)
    printer_2 <- get_stats(second_stats_group, second_stats_name)

    if (!is.null(third_stats_group)){
      printer_3 <- get_stats(third_stats_group, third_stats_name)
      long_note <- paste(printer_1, "\n", printer_2, "\n", printer_3)
    } else {
      long_note <- paste(printer_1, "\n", printer_2)
    }

    dummy_p <- generic_glm_graph(fulldataset, resp_vari, predictor,
"response", pred_name, group_name=folder_tag,

```

```

        fill_col = group_col, save_yesno =
save_yesno, save_name = save_name, note = paste(short_note, group_col, ",
"),
        long_note = long_note)

    print(paste(resp_vari, ": Graph completed.))
    if (print_yesno){
        print(dummy_p)

        writeLines(paste(first_stats_name, " :", resp_vari, "vs. ",
predictor))
        writeLines(printer_r)
        writeLines("\n")
        writeLines(paste(second_stats_name, " :", resp_vari, "vs. ",
predictor))
        writeLines(printer_s)
        writeLines("\nNEW VARIABLES")
    }
}

print("FINAL save_name and GRAPH:")
print(save_name)
return(dummy_p)
}

#EXAMPLES
# generic_grouping_glm_plots_ALL(predictor =
"grad_avg_ratio_rural_suburban_log",
#                               pred_name= "avg rural suburban log",
#                               first_stats_group =
combo_df_zone_onlyrural_dropt,
#                               second_stats_group =
combo_df_zone_onlysuburban_dropt,
#                               first_stats_name = "onlyrural_dropt",
#                               second_stats_name = "onlysuburban_dropt",
#                               fulldataset = combo_df_zone_droptrelease,
#                               third_stats_group = NULL,
#                               third_stats_name = NULL,
#                               short_note = "DropTr,",
#                               group_col = "land_class_binary_l3_site",
#                               print_yesno=TRUE,
#                               save_yesno=TRUE,
#
#                               folder_tag="grad_avg_ratio_rural_suburban_log",
#                               main_dir=mainDir)

# CHANGELOG-----
#V1.6 Apr 23 2025: This includes all the graphs and is the final update of
this
# document probably and I wanted to have a backup.
#
#V1.5 Apr 4 2025: This change includes graphs. That's mostly it. But also
# updates the functionality of other functions to be more user friendly.
#
#V1.4 Apr 1 2025: This change includes the ability to have sex with your
mother.
# It also produces graphs that include the proper statistics for grouping
data
# together. (generic grouping)
#

```

```
#V1.3 Mar 28 2025: Included functionality to create a categorical variable
for
# Age of restoration.
#
#V1.2 Mar 25 2025: Included functions to create ALL THE GRAPHS YOU COULD
EVER NEED
# This version lets you run all the relevant graphs and information for
each
# predictor variable, and also all predictor variables.
#
#V1.1 Mar 25 2025: Included functions to create graphs quickly and easily
that also
# Display relevant information about the p value, test values, test type,
etc.
#
#V1.0 Mar. 20 2025 The first "release" that includes the functions to test
for normality,
# compare continuous and categorical data automatically, and produce a
spreadsheet
# with all possible data combinations.
#
#
```